A SYSTEMS APPROACH TO STUDYING THE DIFFERENTIATION OF STEM CELLS TOWARDS HEPATOCYTES

A Dissertation

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

DAVID CHAU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Adviser: Wei-Shou Hu

August 2017

Acknowledgements

First, I would like to express the deepest gratitude to my advisor, Dr. Wei-Shou Hu, for his invaluable support, guidance, and patience throughout my graduate career. I would not be where I am today without the opportunity that Wei-Shou Hu had given to me. I feel very grateful to have been able to learn from and work with Dr. Wei-Shou Hu who will continue to shape and motivate me into the scientist that I strive to be. His mentorship, wisdom, and lessons will always stay with me as I embark in the next chapter of my scientific career.

I also want to thank Dr. Catherine M. Verfaillie to have been able to work with such a talented researcher in her respective field of stem cell biology.

I would like to acknowledge my thesis committee, Dr. Brenda Ogle and Dr. Wei-Shen in the Department of Biomedical Engineering, and Dr. Samira Azarin in the Department of Chemical Engineering and Materials Science for their patience and commitment to serving on my committee.

I want to specially thank the past and present members of the Hu research group: Jason Owens, Dong Seong Cho, Ravali Raju, Haiyun Pei, Hsu-Yuan Fu, Kyoungho Lee, Liang Zhao, Arpan Bandyopadhyay, Tung Le, Yonsil Park, Andrew Yongky, Kathryn Johnson, Nandita Vishwanathan, Meghan McCann, Sofie O'Brien, Conor O'Brien, Kevin Ortiz, Jennifer One, Thu Phan, Zion Lee, Hansol Kim, Mingyong Xiong, Christopher Stach, Lukas Rositzka, Kristen Cox, and NaTalie Neidenfeuhr for their continued support and encouragement throughout the years of my graduate career. All of you have contributed to enriching my graduate school experience with great memories and I am very grateful for that.

I am so grateful to have met so many good friends in Minnesota that have been the pillars of support throughout my graduate career. I want to specially thank my friends: Avijit Barik, Xinran Lu, Joe Xiao, and Zaw Win, that have also embarked on this journey with me. This journey would not have been the same without your friendship and support.

Lastly and most important of all, I want to thank my wonderful family for their love and never ending support in me. I owe all my success and achievements to their sacrifices. This doctorate degree would not have been possible without having them a part of this journey. To my loving Kirsten Ruud, I cannot thank you enough for being there every step

of the way through the ups and downs of this journey. I could not have imagined having a better partner and companion to have with me during these past few years.

DEDICATION

To my parents Huang Wei and Tuong Han Chau. Without the sacrifices that you made, I would not be where I am today. Without the love and support that you gave me, I would not be who I am today. This thesis is a product of your sacrifices and hard work that you instilled in me.

PERMISSION TO REPRODUCE PUBLISHED MATERIAL

Authorization was granted by the publishers to reproduce the content contained within this dissertation, published during the course of my graduate studies.

1. Raju, R., Chau, D., Verfaillie, C. M., & Hu, W. S. (2013). The road to regenerative liver therapies: the triumphs, trials and tribulations. *Biotechnology advances*, *31*(7), 1085-1093.
2. Raju, R.*, Chau, D.*, Cho, D. S., Park, Y., Verfaillie, C. M., & Hu, W. S. (2017). Cell Expansion During Directed Differentiation of Stem Cells Toward the Hepatic Lineage. *Stem cells and development*, *26*(4), 274-284.

Abstract

The recent advancements in stem cell biology have allowed for new and exciting opportunities to use stem cells in clinical and industrial applications. Stem cells have the unique ability to self-renew and differentiate into any specialized cell type found in the body. Using certain mechanical and biochemical cues, stem cells can be directed to become any specific cell type, such as hepatocytes. A robust and efficient process for expansion and differentiation to generate large quantities of functional hepatocytes from stem cells will be essential to establishing a stem cell bioprocess in the future for therapeutic and industrial applications of hepatocytes.

In this study, a differentiation protocol with soluble growth factors and cytokines was used to mimic the key signaling cues during embryonic development. However, most directed differentiation processes have run into issues with limited scalability and lack of functionality in the differentiated cells. In an effort to bring stem cell therapy closer to reality, our strategy was to use a systems-based approach to enhance the quality and yield of stem cell-derived hepatocytes.

To achieve higher cell yield, we modified an existing differentiation protocol to incorporate a cell expansion stage to facilitate simultaneous differentiation and cell growth. Using transcriptome analysis and mass cytometry, we showed how the population of cells changed over time on both the transcript and protein level. Both analyses revealed that with the new expansion stage, we obtained a higher quantity of hepatocytes within the same time frame compared to the conventional method of differentiation. We then showed the capability to scale up our differentiation for larger scale cultures by adapting the expansion stage onto Cytodex 3 microcarriers. Using the same culture volume as a tissue plate culture, we demonstrated the ability to achieve up to a 5-fold increase in cell number with a final cell density in the range of $4\text{-}5\text{x}10^6$ cells/ml. These strategies show that the demand for large quantities of hepatocytes can be met by translating the conventional method of differentiation to suspension microcarrier differentiation.

Encouraged by our ability to yield higher cell density using microcarrier culture, we explored assessing the functional maturity of our stem cell-derived hepatocytes using transcriptome analysis. We showed that stem cell-derived hepatocytes are still clearly different when compared to primary hepatocytes at the transcriptome level. In addition to

evaluating cells using transcriptome analysis, we wanted to be able to compare the current *in-vitro* processes to embryonic liver development to understand the genetic roadblocks. The transcriptome data from hESCs hepatocyte differentiation was integrated with mouse liver development using principal component analysis and batch corrections. This allowed us to create a unified developmental scale to compare samples from different species and *in-vitro* to *in-vivo* platforms. The meta-analysis revealed that stem cell-derived hepatocytes are equivalent to the functional maturity of developing cells at E15 in mouse development.

From the transcriptome analysis, we observed many different genes in energy metabolism with dynamic behavior over the course of differentiation. We sought to understand the effect of changes in different metabolic genes and the impacts on metabolic transition during differentiation. We characterized the energy metabolism of hESCs and assessed the metabolic demand of cells at different stages of differentiation. hESCs and early differentiated cells exhibited a high glycolytic flux. transitioning towards an oxidative metabolism as the differentiation progressed. Furthermore, using confocal microscopy, we also characterized the activity and morphology of the mitochondria in the cells at different stages of differentiation. Using the consumption rates of different nutrients as an input to our metabolic flux model along with our transcriptome findings, we were able to gain a deeper understanding of the metabolic behavior of cells during differentiation. Our analysis revealed that cells consume lower amounts of glucose over the course of the differentiation but become more efficient at transporting pyruvate into the mitochondria leading to increased oxidative phosphorylation. However, our metabolic and transcriptome data revealed that our stem cell-derived hepatocytes are not capable of mature metabolic functions such as gluconeogenesis, supporting the immature phenotype that has been described in literature.

Together, these studies reveal that stem cells can provide a renewable and scalable source of hepatocytes for therapeutic applications. These cells demonstrate some phenotypic and functional properties of primary hepatocytes but have some contrasting elements compared to their *in-vivo* counterparts that will need further genetic intervention to enhance their maturation before cellular therapy can become a reality. However, this work is invaluable as it contributes to the current status of the field and facilitates the translation of laboratory practices of stem cell culture into a scalable technology.

Table of Contents

List of Figures

List of Tables

Chapter 1 : Introduction

1.1 Introduction

The liver is the largest internal organ inside the human body. This organ is responsible for many important functions including drug metabolism and detoxification, protein synthesis, energy metabolism, synthesis of biochemical for digestions, and many more functions that are necessary for maintaining homeostasis within the organism. Many of these functions are carried out by the parenchymal cells found in the liver, commonly referred to as hepatocytes, which make up 70-85% of the liver mass.

Although the liver is inherently capable of regeneration following any acute injury, typical liver disorders and end-stage liver diseases can compromise the regenerative capability and result in liver failure and ultimately death. Currently, there are many options to alleviate the effects of liver failure. However, there is only one viable treatment option through liver transplantation. As of 2016, there are currently over 14,000 people on the waitlist for liver transplantation in the United States alone and half of those patients will die annually due to the shortage of donors (1). Thus, the shortage of donor organs and risks present with liver transplantation presents serious limitations to the liver transplantation approach that will need to be addressed.

There have been alternative solutions that have been proposed to address these issues. Hepatocyte transplantation has shown to have promising results especially in liver-based congenital and metabolic disorders. Bioartificial liver devices which rely on using hepatocytes from a variety of cell source (e.g. porcine, human, stem cell derived, etc.) can use the mechanical and cellular components of the device to provide temporary assistance until a donor organ is available. In these alternative treatment options, primary human hepatocytes are the preferred cell source due to its lower risk in immunogenicity and exact match with the human physiology. However, primary hepatocytes have limited proliferation and the availability of donor liver for isolation of liver cells is still far below what is demanded.

With the recent advances in developmental biology and the emergence of stem cell therapy, many have attempted to use stem cells as an alternative cell source for obtaining

hepatocytes. Using our understanding of embryonic liver development, many groups including our own have developed a differentiation protocol to use different biochemical and/or mechanical cues to mimic the embryonic development for generating hepatocytes from stem cells. Through the sequential stages of different growth factors and cytokines, stem cells can be guided through different developmental stages to give rise to cells, referred to as hepatocyte-like cells, that are functionally and phenotypically similar to primary hepatocytes.

Stem cell fate determination is typically the result of very complex processes through the combination of soluble signals, physical stimuli, and cell-cell interaction. Embryonic stem cells are isolated from the inner cell mass during embryonic development. It can be hypothesized then that the differentiation of stem cells towards the hepatic lineage will follow similar progression as *in-vivo* liver development. In recent history, stem cell research has made tremendous strides in using different strategies to generating hepatocytes from different stem cell and progenitor cell populations. The simplicity of the culture systems developed for *in-vitro* differentiation will need to be well thought out to meet the demands in the complexity of embryonic development. In addition, the issue of being able to design different platforms for generating hepatocyte cells in a scalable manner to meet the demand for hepatocytes in clinical and industrial applications will also have to be kept in mind.

1.2 Research Objectives

The goal of this study is to establish a robust platform for the generation of functional and mature stem cell-derived hepatocytes in a scalable manner for clinical and industrial applications. The key objectives in meeting this goal are:

Objective 1: To improve the cell yield of the directed differentiation of hESCs to hepatocytes by incorporating an endodermal expansion stage to mimic the embryonic development

Objective 2: To develop a microcarrier based culture system for scalable expansion and differentiation of endodermal cells towards the hepatic lineage

Objective 3: To perform a systems-based study using transcriptome data to understand the transcriptional control of different gene expression that are vital to generating hepatocyte-like cells

Objective 4: To understand the metabolic changes that hESCs undergo during the differentiation to hepatocyte-like cells and characterize the nutrients demand of cells at different stages of differentiation for future studies using environmental monitoring and control

1.3 Scope of Thesis

A detailed literature review is presented in Chapter 2. The first part of this chapter serves as an introduction to using stem cells and its defining characteristics to address liver failure. The second part of the chapter provides background information on our current understanding in the metabolic profile of stem cells and hepatocytes. This has been utilized to gauge the progression of stem cells towards the hepatic lineage and identify the ability of stem cells to adopt the metabolic elements of a hepatocyte. The final part of this chapter highlights the current status in in stem cell technologies and the options that are available for generating hepatocytes from stem cells in a bioprocess manner.

Chapter 3 presents a detailed list of all the materials and universal methods that are employed in the different studies in this thesis. Chapter 4 focuses how cell yield can be improved by utilizing an endodermal cell population and comparing it to the conventional method of differentiation at the transcriptomic and protein level. Chapter 5 presents a comprehensive study between in-vitro differentiation and in-vivo development to understand the progression of stem cells towards the hepatic lineage and what further steps need to be taken to improve the maturity of stem cell-derived hepatocytes. Chapter 6 focuses on understanding the metabolic profile of our cells during the differentiation process and how to use different cell culture conditions to improve the maturity of our differentiation for future scalable reactor studies. We also describe the studies that were conducted to improve the scalability of differentiation by adapting our protocol described in chapter 4 to being carried out on microcarriers. Finally, in Chapter 7, a conclusion summarizing the findings and their relevance to the field of stem cell is presented along with potential routes for the future work that still need to be investigated further.

Chapter 2. Background

2.1 An Introduction to Stem Cells

Stem cells are a unique cell type defined by their ability to self-renew, to differentiate into different cell types, and to functionally reconstitute tissue *in-vivo*. These capabilities make stem cells an attractive candidate in the field of regenerative medicine and tissue engineering. With the emergence of stem cells and as our understanding on stem cells continue to grow, we can begin asking fundamental questions on the mechanisms behind the self-renewal and different differentiation processes to become specialized cell types that make up our body. Currently, the field has already begun to evaluate the potential of stem cells in different applications for studying different diseases and drug toxicity studies. As our knowledge in stem cell biology improves, we can begin to evaluate the possibility of using stem cell based therapy to treating diseases like diabetes, liver cirrhosis, Parkinson's disease, and many others.

Stem cell technology will play a key part in the utilization of stem cells not only for the purposes of research and fundamental biology but also for the pharmaceutical industry in therapeutical settings as well as for drug screening studies. The hopes of the next revolution of medicine with stem cells will rely heavily on our understanding of the biology and mechanisms regulating stem cells and its properties. In these next few sections, we will be discussing the different properties of stem cells and the impact it can have for regenerative medicine especially for liver failure. Finally, we will explore the status of generating hepatocytes from different types of stem cells and the properties of those stem cell-derived hepatocytes.

2.2 Properties of Stem Cells

A cell is characterized as a stem cell for fulfilling three unique characteristics: self-renewal, potential for multi-lineage differentiation, and *in-vivo* reconstitution of certain tissues.

2.2.1 Self-Renewal

Self-renewal is typically the first property that is used to define a stem cell. Normal somatic cells can only undergo a limited number of cell division before senescence occurs. This phenomenon is typically referred to as the Hayflick's limit (2). Unlike somatic cells however, stem cells have the capability to self-renew in its potency

state through large numbers of cell division giving rise to more daughter cells with the same genetic properties as the parental cell. Although the mechanisms of self-renewal still remain a major topic of research, it has long been accepted that the immortal proliferative capacity of tumor and cancerous cells have been linked to the activity of the enzyme, telomerase, and maintenance of the telomere length (3, 4). Thus, there have been many studies and debates on the activation of the telomerase in playing a role for the self-renewal capability in stem cells. Recently, there is growing evidence that telomerase is activated in different stem cells and the level of activation is dependent on the type of stem cells which gives rise to the different rates at which senescence occur at (5). It should be noted that the activation of telomerase to relieve telomere shortening is a necessary step for the onset of cancer and have resulted in heavy investigations between the mechanisms between self-renewal and the mechanisms of cancer. There are still ongoing debates as to whether the activation of telomerase and slowing down of telomeres will be the key to longevity but how that is balanced with the onset of tumor remains to be seen. Through these efforts, a different type of stem cells is believed to exist within different tumors and cancer that is now commonly referred to as "cancer" stem cells. These cancer stem cells are believed to be the reason for metastasis of cancer even after different cancer treatments but the existence of these cells are still under heavy debate that is beyond the scope of this thesis. Nonetheless, the reasons behind the self-renewal behavior of stem cells and the factors regulating this process will be critical for the purposes of obtaining large numbers of cells in a renewable manner for the different clinical and industrial applications envisioned for regenerative medicine and tissue engineering (6).

2.2.2 Potency

The second property that stem cells possess that will be described in this chapter is the ability of stem cells to differentiate into specialized cell types that make up the human body. The differentiation potential of a stem cell is defined as its potency. The regulation of potency and self-renewal typically go hand in hand and recent studies have shown that the mechanisms behind the degree of potency and self-renewal might be more closely related than previously thought. Stem cells can remain uncommitted for long periods of time generating new daughter cells with the same potency. It is because of these behaviors

that stem cells has become a key renewable cell source for regenerative medicine. Several transcription factors have been found to be important regulators in maintaining the undifferentiated state of a stem cell along with different chromatin-modifying enzymes, regulatory RNA molecules, and signal-transduction pathways (7-11). Three major transcription factors (Oct4, Sox2, Nanog) have been found to be essential in promoting the expression of genes to maintain pluripotency while also repressing genes that induce differentiation (12). The importance of these transcription factors was further demonstrated when somatic cells were transfected with these factors were found to become induced pluripotent stem cells (iPSCs).

There are different degrees of potency that is generally designated based on the number of cell types that the stem cell can differentiate into. Typically, potency can be divided into four major classes: totipotency, pluripotency, multipotency, and unipotency. Totipotent cells can become any cell type of embryonic and extraembryonic origin found in the organism (13). These cells are typically found after the fusion between a sperm cell and an egg cell. Pluripotent stem cells are stem cells designed as capable of differentiating into any of the three germ layers: endoderm, ectoderm, and endoderm (13) . The most commonly found pluripotent stem cells are embryonic stem cells or iPSCs. Multipotent stem cells are typically capable of differentiating into a limited number of cell types typically to a subset of cells found such as cells found within the hematopoietic system (13). Unipotent stem cells are the least potent type of stem cells that can give rise to only one particular specialized type but is still capable of self-renewal. These types of cells are typically referred to as precursor cells such as hepatic precursor cells or cardiac precursor cells (14, 15).

2.3 Types of Stem Cells

2.3.1 Pluripotent Stem Cells

Embryonic stem cells (ESCs) are pluripotent stem cells derived from the inner cell mass during the blastocyst stage of development. These type of stem cells are still the most widely studied in stem cell research. Mouse ESCs were first isolated in 1981 and were found to have the remarkable capability to re-enter embryogenesis after injection into a pre-implantation embryo and functionally contribute to all the tissues and organs of a new mouse (16). Due to the pluripotent nature of cells, they emerged as promising potential

tools for clinical applications. These isolated mouse ESCs were then injected into animal models and pre-implantation embryo and found capable of forming teratomas consisting of tissues from all three germ layers. Then, in 1998, Thomson et al. was able to successful derive human ESCs from human blastocysts showing remarkably similar properties as mouse ESCs. Both human and mouse ESCS were further characterized by the expression of the transcription factors, Oct4, Nanog, Sox2, Rex, and the surface markers SSEA-1 (17). The expression profile of pluripotent cells were universally defined by the presence and expression of these key genes that are now commonly referred to as pluripotent genes. In addition to the phenotypic expression profile, ESCs were found to have unlimited proliferative capacity along with a high nucleus to cytoplasm ratio (18, 19). These cells were initially isolated and routinely cultured on a layer of irradiated mouse embryonic fibroblasts (MEFs) in the presence of serum. The MEFs were found to release certain growth factors and cytokines that help maintain the pluripotent nature of ESCs (19, 20). Although, for the purposes of clinical and industrial applications, there will need to be extensive studies to try and replace MEFs and serum to have more defined media components either using various cytokines or other sources of feeder layer from human (21). In addition, the research of ESCs had a period of controversy if it was ethical to conduct research that involved the development, use, and destruction of human embryos. In addition, there was issues raised whether ESCs-derived cells for cellular therapy would trigger an immune response since ESCs were derived from embryos with a specific immunogenic state and not designed to specifically match the immunogenic state of the patient (22-25). Nonetheless, as the stem cell and regenerative medicine field begin to advance, studies have shown that ESCs can be used to treat certain disorders and diseases in numerous clinical trials.

However, at the time due to these concerns, in 2006 a research group in Japan discovered a new pluripotent population called induced pluripotent stem cells(12). These cells showed the pluripotent nature but was generated through a process termed “cellular reprogramming”. Through the expression of a combination of different key transcription factors involved in pluripotency, somatic cells could be reprogrammed to a stem cell-like fate and exhibit many pluripotent characteristics simply through the transfection of Oct4, SOX2, KLF4, and C-myc. These factors are now known as the Yamanaka factors to credit

the scientist that had discovered them. Since iPSCs can be derived from any somatic cells isolated from an individual, they not only bypass the need for embryos but it can be thought of as personalized medicine by using an autologous cell source that should not have the risk of immune rejection. To demonstrate the therapeutic concept of using iPSCs, a group used fibroblasts from a humanized sickle cell anemia model mice and reprogrammed them into iPSCs. The diseases iPSCs were corrected by using gene targeting and differentiated to hematopoietic progenitor cells to be transplanted back into the diseased mice showing rescue of the mice with the corrected iPSC-derived hematopoietic cells (26).

Pluripotent stem cells can be directed to differentiate towards any lineage of the three germ layers. There are a number of differentiation strategies that utilizes different mechanical or biochemical cues to mimic the embryonic development. One strategy is to allow stem cells to spontaneously differentiate and select for the desired cell type by sorting or selecting those that have the desired phenotype. However, this method can result in very heterogeneous culture with typically very low yield in desired cell type along with tedious labor to obtain the cell of interest. Thus, more recently as our understanding in developmental biology continues to improve, strategies of using directed differentiation through the addition of various cytokines and growth factors to mimic the signaling of in-vivo development along with co-culture with certain cell types found in development can achieve in much higher cell yield and cell quality. Thus, while ESCs have the versatility to be used in a number of different applications, there are still many obstacles and unknowns that need to be sorted out before clinical therapies in human can be reached.

2.3.2 Adult Stem Cells

Adult stem cells are typically derived from the tissues of post-natal or adult animals. These cells typically have a lower potency and limited self-renewal capability compared to pluripotent stem cells. The most studied adult stem cells to date have been isolated from the bone marrow and are typically involved in the hematopoietic system. Adult stem cells have been believed to exist in every organ but exist at very low numbers and only resides in a stem cell ‘niche’ that favor’s their quiescent state. Once activated, they can proliferate and give rise to more adult stem cells and daughter cells that will differentiate towards a cell lineage. Although these properties are still being studied today, it is intuitive that these cells remain in an inactive state as they have limited proliferative capability but become

activated to replenish the cell population after the native population undergoes senescence. Once activated, these adult stem cells can proliferate and their daughter cells can either remain as stem cells or differentiate to the other specialize cell types. There are a number of different adult stem cells found in different organs and niches. Hematopoietic stem cells found in the bone marrow are the most well-known and well-studied adult stem cells. Neural stem cells are isolated from the brain and can give rise to the different populations that are found in the brain. Mesenchymal stem cells are of the stromal origin that can be isolated from many of different tissues including lung, bone marrow, adipose, and molar.

2.3.2.1 Liver Stem Cells

In our study, we were interested in generating hepatocytes from embryonic stem cells for the purposes of different liver applications. Thus, we considered the different populations of liver stem cells. Hepatocytes in the adult liver are capable of inherent regeneration upon acute liver injury. The isolation of hepatocytes from an adult mouse model have shown to be capable of up to 70 doublings and still engraft in the liver mass through serial isolation and transplantation (27). However, in certain incidents where there is chronic liver failure or severe liver injury, the regenerative capability of hepatocytes become compromise and the resident hepatic stem cells can proliferate and repair the liver in a process known as ductular reaction (28). There are still many debates and questions that remain as to how this process occurs and the mechanisms that trigger this process. However, many researchers and clinicians have referred to the proliferative hepatic stem cells that participates in regeneration as oval stem cells.

Both populations of hepatocytes and the hepatic stem cells in the adult liver are capable of proliferation and reconstitution of the liver mass. Such potential can potentially be harnessed for in vitro culture and clinical applications. However, our current culture systems have resulted in the lost of proliferative capacity as well as functional activity for hepatocytes. With the advances in our understanding of the nature of liver stem/progenitor cells, our ability to isolate and culture liver stem cells or progenitor cells has enhanced greatly in the past few years and will be key in translating this knowledge to differentiating hESCs into hepatocytes.

Study of liver stem/progenitor cells often employs partial hepatectomy or treatment with drugs like acetylaminofluoren in rodents. By BrdU labeling of proliferative cells in regenerating livers of rodents after induced acetaminophen injury, the region that stem cells reside and their niches were identified. Four differently labeled stem cell populations were identified, namely periductal mononuclear cells and peribiliary hepatocytes along with cells in the canal of Hering and intralobular bile ducts (29). While most have accepted hepatoblasts as a key contributor to liver regeneration, Kuwahara's finding raises the interesting notion that there may be multiple sources of this cell source that may be responsible for liver regeneration.

The availability of genomic data and the enhanced capability to identify genes differentially expressed under different differentiation states helped identified markers suitable for cell isolation. The markers identified in rodents enabled selective isolation of the rodent counter parts in human liver and revealed interesting similarity between embryo liver development and adult liver regeneration in terms of delineation of liver cell differentiation (30-32).

EPCAM has been a key marker for enriching both fetal and adult liver stem/progenitor cells in rodents as well as human(33-36). Recently, the detection of high level expression of aldehyde dehydrogenase using a fluorescent substrate, adult liver stem cells from both mice and human liver were isolated without using any chemical stimuli to enrich their population (32). This methodology could possibly provide a readily robust method of isolating a relatively homogenous population of hepatic progenitor cells. A recent review by Turner et al. summarizes current state of understanding of the locale of liver stem cells in liver lobule, their characteristics in human (37). Readers are also referred to a more detail overview of the fetal and adults liver stem cells (38).

Turner et al. presented a comprehensive descriptive progression of liver stem cell during the development of the liver. It is postulated that the hepatic stem cells reside in the periportal region of the liver, specifically the canal of Hering. As these cells proliferate "older" cells migrate along the cell plate toward central vein and grow more committed and mature; first becoming bipotential, then becoming proliferative hepatocytes, and mature hepatocyte. As they move toward central vein, they also become larger in size.

Further details regarding the origin and repair mechanisms are described elsewhere (37-39)

The hepatic stem cells have been isolated from both fetal liver and adult liver; whereas the mesonendodermal cells have only been reported from human fetal liver. Both precede the hepatoblast; although both share many markers of hepatoblasts including EPCAM, E-cadherin, CK8, CK18, CK19, both are also AFP and albumin negative. A distinction between hepatic stem cells and hepatoblast cells is their NCAM/ICAM-1 expression; ++/- and -/++ for hepatic stem cell and hepatoblast respectively.

Hepatic stem cells isolated from human adult liver are about 7-10 μm in diameter (32, 37); somewhat smaller than hepatoblasts in both adult and fetal liver (10-12 μm). They exhibit a high nucleus to cytoplasm ratio. Hepatoblasts are abundant in fetal liver (~80%), but decreases after birth to less than 0.01% in adult liver. In contrast, hepatic stem cells remain relatively stable comprising of about 0.5 to 1.5% of the liver cell mass throughout (40).

In addition to growth factors, surface properties also affect the outcome of cell fate. Turner et al. reported that the both hepatic stem cells and hepatoblasts grew out from EPCAM sorted cell population when plated on plastic surface, but only hepatic stem cell colonies emerge when plated on STO feeder layer or a surface treated with ECMs like collagen I collagen III or angioblast feeder layer for hepatic stem cells and collagen IV, laminin or stromal feeder cells for hepatoblasts (30, 31, 33).

2.4 Differentiation of Stem Cells to Hepatic Lineage

Because of the liver's high regenerative capacity there has been a long history of research on liver stem/progenitor cells and on their in vitro expansion. Those attempts aim to expand cells already committed to a developmental path. This contrasts with the effort initiated after the isolation of embryonic stem cells, which seek to direct uncommitted cells to endoderm lineage and further down to hepatic path. Embryonic stem cells have huge self-renewal and differentiation capacity. If they can be expanded and guided to liver lineage a potentially unlimited supply of hepatocytes may ensue. The emergence of iPSCs only amplified the effort.

2.4.1 Differentiation of Hepatic Stem Cells

Hepatoblasts and liver stem cells isolated from fetal and adult livers have been shown to differentiate to hepatocyte-like cells in culture. A combination of growth factors, HGF, EGF, Oncostatin and Dexamethasone administered over one week was used to induce hepatic differentiation for mouse hepatoblasts (32). Recently it was shown that the liver progenitor cells isolated from human adults can be expanded and differentiated to hepatocyte-like cells with many liver functions including urea synthesis, albumin secretion and CYP activity. Human hepatic stem cells were expanded in a medium known as Kubota's medium which is a serum free medium composition with no copper and low calcium. These cells were induced to differentiate by supplementing Kubota's medium with calcium, copper, glucagon and tri-iodothyronine. In this study, decellularized liver biomatrix was used as an ECM thereby reducing the need of commonly used growth factors HGF and EGF (41).

The development of liver in vivo entails the specification to definitive endoderm (about E8.5 in mouse), followed by commitment to hepatoblast and formation of liver bud, and finally fully differentiated hepatocytes and the emergence of bile ducts. The progression of the development is guided by a number of inductive signals dynamically. The early differentiation to mesoendoderm and subsequent distinction of mesenchyme and endoderm is driven by Nodal ,BMPs, and Activin signaling (42-44). Further signaling from the FGF and BMP family, specifically BMP4, FGF1, FGF2,FGF4 and FGF8, induce differentiation to hepatoblast. Liver bud formation is guided by the inductive signal of HGF and oncostatin which stimulate the hepatoblasts to differentiate towards hepatocytes.

Most efforts in guiding the differentiation of pluripotent stem cells to hepatic lineage deploy growth factors to mimic the temporal dynamics of cues in hepatogenesis in-vivo. However, it should be noted that the development of the liver and any other tissue in vivo is a continuous process, whereas our efforts to replicate these processes tends to be in discrete stages. Each various protocol for guiding the differentiation of stem cells to hepatic lineage are segmented into different differentiation stages. Thus, each protocol may implement different numbers of stages, and different corresponding duration and medium composition. Representative protocols used for differentiation to hepatic lineage in vitro are listed in Figure 2-1.

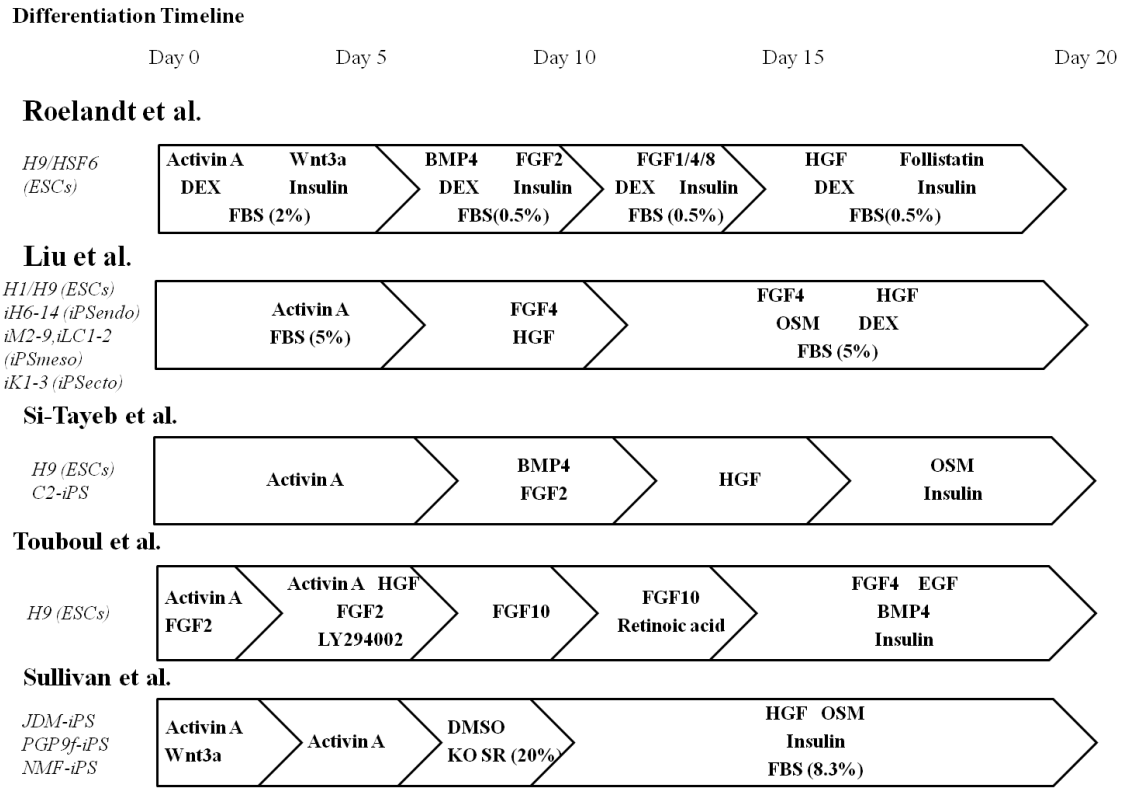


Figure 2-1. Hepatic Differentiation Protocols for iPSCs and ESCs.

Most hepatocyte differentiation protocols start with Activin and Wnt3a for a 3-5 day period to induce definitive endoderm commitment (42, 45-47). FGF and BMP plays a significant inductive role in promoting endodermal progenitor maintenance and expansion during the second stage of directed differentiation (48). Further treatment with a mixture of FGFs facilitates the commitment to hepatic fate or the equivalent of hepatoblast state in liver development. The last stage often entails the use of oncostatin or follistatin and hepatocyte growth factor (HGF). It was previously reported that oncostatin-treated stem cells would lead to an increased activity level of STAT3 (49). HGF, shown to contribute to liver development in a STAT3-independent manner, is also often used in the last stage of directed differentiation toward hepatic lineage (49).

2.4.2 Differentiation of Stem Cells using Gene Transfection

A number of laboratories have demonstrated the differentiation of iPSCs to hepatocyte-like cells using the protocols developed for ES cells. It appears that the protocols developed for embryonic stem cells are mostly applicable to iPSCs. This opens the possibility of personalized hepatocyte-like cells for drug screening and even for

therapy. However, deriving iPSC is a lengthy process. Recently the feasibility of directly inducing fibroblastic cells to hepatic lineage through gene transfection mediated reprogramming was demonstrated in mice models. A combination of Hnf4a with Foxa1, Foxa2 or Foxa3 was used to transfect and reprogram adult mouse fibroblasts to hepatocytes-like cells which are called iHEP cells (50). The efficiency was about 1 in 1000 of initially transfected embryonic fibroblasts. Interestingly the resulting iHep cells possess a very high proliferative potential, unlike the hepatocyte-like cells derived from from pluripotent cells which appear to die off in prolonged culture. In another study the over expression of Gata4, Hnf1α and Foxa3 along with $p19^{Arf}$ inactivation enabled mouse fibroblasts to become iHep cells (51). In this case inactivation of $p19^{Arf}$ suppressed senescence to enhance the chance of isolating reprogrammed cells.

The reprogrammed iHep cells expressed key hepatic genes and other liver functions and successful engraftment was shown in animal models. The obvious question then is why does the resulting iHep cells have a higher proliferative potential compared to the hepatocyte-like cells derived from pluripotent cells. The differences and similarities between the different types of hepatocyte-like cells derived from stem cells along with primary hepatic stem cells have yet to be studied extensively. However, the approach will certainly see many in vitro applications if it can be shown to work in reprogramming human cells.

2.5 Current Technology of Stem Cell Culture

With the recent advancements in stem cell biology and potential of stem cells for different applications, there has been an increasing interest in the development of stem cell bioprocess. In this area, researchers are focused on developing new ways for scalable expansion and differentiation of stem cells to meet the demand of different cells for different applications. However, current stem cell practices are routinely carried out on tissue culture plates where it does not come close to the scale that would be needed for tissue engineering and regenerative medicine. In addition, the need to have a robust bioprocess to have consistent and reproducible processes is very difficult to achieve in a tissue culture plate. As stem cells is very susceptible to slight changes in different culture environments, the inability to monitor and control the culture conditions poses a severe

problem to achieving well-defined outcomes. Thus, there has been large focus on using different bioreactor configurations to address the demand for large quantities of cells and have well-defined culture conditions to scale-up cultures with the ability to achieve continuous monitoring and control of the microenvironment of the cells. In this section, we will be focused on the recent studies that has been carried out on translating laboratory stem cell practices to a technology.

The most common cell culture practice carried out today is used for the manufacturing of different therapeutic proteins in the biopharmaceutical industry. For these biological processes, the end goal is to use a variety of cells (e.g. insect cells, Chinese Hamster Ovary Cells) to secrete the protein of interest in large tank reactors in a scalable manner. But as we discussed earlier before, the fundamental difference between future stem cell processes and current biological processes is that future stem cell bioprocess will require the separation and purification of the cells themselves from the process. To this end, we are interested in the generation and harvesting of the cells itself which will be our product. Thus, when designing a cell culture process for manufacturing stem cells we need to keep in mind what our product and objective is.

2.5.1 Reactor Configurations for Stem Cell Culture

Typical cell culture processes use different bioreactor systems for cell cultivation for a more scalable process. There are many different bioreactor configurations that are used in the industry to meet the demands of cells and the particular process in mind. However, the fundamental bioprocess engineering principles that include bioreactor design and process control that have been established for the traditional pharmaceutical industry can still be translated to stem cell cultivation.

The most commonly used reactor system used today is the stirred reactor system where an impeller is use to mix the culture to achieve a homogenous environment and effective nutrient delivery. Most reactor systems used today can be coupled with automated programs to allow for continuous monitoring of the environment and different feedback controls to regulate the various chemical and physical parameters in the bioreactor (52, 53) Simple stirred reactors are easy to operate and facilitate larger quantity of cells with access to continuous information of the process run. However, the presence of an impeller to

ensure homogenous conditions also poses hydrodynamic shear stress to the cells. A sufficiently high agitation rate should be established to ensure a homogenous environment and prevent agglomeration and microcarrier bridging without introducing too much shear stress. Although the simplicity and ease of stirred tank bioreactors have pushed it to the forefront of the bioreactor configurations for scalability, there are still many effects that have not been examined and other reactor configurations that can also meet the demand for scalability.

The wave bioreactor is another bioreactor design that have been used in expanding cells especially in the area of cellular therapy for small to medium scales (54). These reactors are typically consisted of a disposable bag that is inoculated with media and cells that can be connected to several probes for monitoring. The wave reactor allows for ease of use and lower risk of contamination because of the pre-sterilization and one-time use. These culture bags can be placed on an apparatus to allow a rocking motion to create waves at the air-liquid interface in the bag, hence the name wave reactors. This will ensure a homogenous environment without the introduction of mechanical agitation seen in stirred tank reactors. However, this configuration has proven to be higher in cost and the demonstration of scalability is still limited.

In the field of bioprocess manufacturing, there has been a recent push toward continuous process with an emphasis of using perfusion systems. Currently, most reactor cultures are done in batch settings and scaled up from the seed train at smaller scales all the way to the manufacturing scale. However, this approach only harvests the product of interest at the largest scale where there are debates whether this is done in the most economical and efficient manner. In a perfusion setting, there would be a continuous process of harvesting and maintenance. Current reactor configurations can be modified to operate in a perfusion culture with continual exchange of media as fresh or recycled media is introduced (55). Currently, there are other bioreactor configurations that are designed specifically to serve a perfusion run; fixed bed, fluidized bed, and hollow fiber.

One of the biggest difficulties in scaling stem cell culture is the ability to adapt stem cells to suspension culture for the different reactor configurations. Typically, most cells

grown in reactor systems can be grown as single cells or aggregates in suspension to allow for scalability. For non-adherent cells, single cells or aggregates are likely to be cultured in stirred tank to prevent the agglomeration and homogenous environment. As mentioned before, stem cells however, face very low viability and spontaneous differentiation when cultured as single cells. There are some exception to certain stem cells such as hematopoietic stem cells which have been demonstrated to be viable as a suspension culture(56). Otherwise, most stem cells are unable to survive as dissociated single cells in a suspension culture (57). Thus, currently there are three cultivation formats that have been used for stem cell processes for scalability: i) non-adherent suspension culture ii) adherent layer culture and iii) aggregate culture. Although most studies with stem cells are in the late phases of research and clinical trials, the current trend points towards the need to develop a platform to establish high cell doses for therapeutic purposes after the clinical trial phase.

For therapies that would require large quantities of cells, a lot of studies have relied on using the traditional stirred tank bioreactors. For instances, a lot of work have been focused on establishing the feasibility of maintaining and expanding mesenchymal stem cells (58-61), pluripotent stem cells (62-66), and hematopoietic stem cells (56, 67) in stirred tank bioreactors. These studies have established the foundation of using bioreactor systems to propagate undifferentiated cells in a scalable manner for inducing differentiation to a large quantities of cells towards a particular lineage. Most of these studies use microcarriers to provide a solid substrate for stem cells to adhere to for growth and survival. Microcarriers have long been used in the biopharmaceutical industry especially in the field of virus production (68-72). Because of the demonstration and understanding through current processes that use microcarriers in large scale settings, we believe that microcarriers are a viable option for scaling stem cell culture practices. The success of using microcarriers for stem cell culture practices will depend largely on the separation process which still pose as a critical issue for stem cell bioprocessing. There has been extensive research that are still ongoing regarding the design for isolation and purification that can be found in this comprehensive review (73).

In addition to the different configurations of bioreactor systems, there are a number of different microcarriers that are out on the market that can be used for stem cell processes. The most commonly used microcarriers are solid microcarriers that are made of cross-linked dextran, cellulose or polystyrene. These microcarriers allow for cell attachment and growth on the surface of the beads. Macroporous microcarriers are another type of microcarriers that are made of either gelatin or collagen that have pores that reach into the beads to provide additional surface area for further attachment and growth.

Table 2-1. Attachment Efficiency of Human Embryonic Stem Cells on Different Microcarriers. [Modified from Thakur et al. 2015]

Microcarrier	Material	Cell Attachment	Cell Viability
Cytodex 1	Cross-linked dextran with N,N diethylaminoethyl groups	Fair	Poor
Cytodex 3	Cross-linked dextran immobilized with denatured collagen	Fair	Fair
Culti-sphere S	Crosslinked pharmaceutical grade porcine gelatin	Fair	Poor
Solo Hill F102	Crosslinked polystyrene modified with cationic gelatin	Poor	N/A
Solo Hill C102	Crosslinked polystyrene modified with gelatin	Poor	N/A
Solo Hill PP102	Cationic crosslinked polystyrene	Poor	N/A

2.5.2 Media Design for Stem Cell Culture

One of the biggest market and influential scientific aspects in the stem cell industry is the media development. The current market for the stem cell industry was valued to be at $8.8 billion US dollars as of 2016 (74). Approximately a third of that value is estimated to come from stem cell products such as media and the different components. As our interest and potential of using stem cell grows, investment into this sector is expected to continue to grow at an exponential rate. Overall, we must harness our understanding of the different pathways and its effect on stem cell behavior in order to move the field forward. Without the proper medium and signaling cues, stem cells cannot behave the way we expect them to. In the past few decades, the focus of media development have been

optimized for the industrial mammalian cell culture processes for biologics. With the growing importance of stem cell research, the focus of medium design has been extended to enhancing stem cell behavior and growth. Most cell types share a basic nutritional requirement to promote survival and growth. However, stem cells are more susceptible to varying levels of nutrients and the need for different growth factors and cytokines increasing the complexity of the design. In order to continue our progress in this sector, we must also improve in our techniques to meet the demand to grow and differentiate stem cells. In this section, we will focus on the different nutritional needs of stem cells that are known and how it will influence the translation of current stem cell research into a technology.

Most optimal medium are designed to mimic the chemical environment of the native environment of the cell origin. Although the native chemical environment has been proven to be feasible to maintain and promote survival of different cells, studies continue to provide us with additional information on the effect of different components on cell physiology and behavior. Depending on the objective of the process, a cell culture medium can vary its composition to meet the demand of the cells. Typical cell culture media contains small molecular weight components that are found in the interstitial fluid of the body. Although, the chemical composition may vary from tissue to tissue, there is a general composition of sugar, amino acids, vitamins, nucleotides, salts, and trace elements that compose the interstitial fluid.

The most common supplemented sugar in cell culture medium is glucose. Glucose is the basic component found in cell culture medium to provide fuel to the cells in the form of high energy phosphate bonds in ATP, hydrogen atoms associated through the co-enzymes NADP and NAD as well as other nucleotide triphosphate (75, 76). There are other forms of simple sugars such as fructose, mannose, and galactose that can provide energy to cells as well. However, the utilization of the sugar is dependent on the presence of the corresponding sugar transporter. GLUT is the transmembrane protein responsible for enabling transport of different sugars into the cell for metabolism. However, the different sugars are uptake and transported depending on which isozyme of GLUT is expressed. In most mammalian cells, GLUT1 is the predominant isozyme. Although

GLUT1 is predominantly responsible for transporting glucose, it has also been shown to transport other sugars such as galactose, mannose, and glucosamine (77). In stem cells where glycolysis is known to promote maintenance and growth, GLUT1 and GLUT3 are known to be the predominant isozyme (78, 79). Both isozymes have a fairly low K_m making it have a higher affinity to glucose compared to the other GLUT transporters (80, 81). As a result, most mammalian cell culture utilize glucose as the major sugar for carbon in its medium. The physiological concentration of glucose in interstitial fluid is approximately 1g/L. As a result, most cell culture medium is present from 1g/L to 5g/L depending on the application. The conventional wisdom is that the presence of higher glucose levels can lead to larger uptake of the glucose facilitating faster growth and lead to higher production of different proteins. However, recent studies have shown that high glucose can have detrimental effects on stem cells and its behavior (82, 83). However, the inherent glycolytic metabolism of stem cells requires that its medium be consist of high levels of glucose to prevent starvation of glucose and meet the demand of its high glycolytic state.

Nucleosides are another basic component present in basal medium that provide the cells with the basic building blocks to RNA and DNA. Nucleosides consists of adenosine, cytidine, guanosine, and uridine. The supplementation of nucleosides either through serum or exogenous nucleosides in serum-free medium can promote cell growth (84). When replication stress occurs in highly proliferative cells such as stem cells, nucleosides can be supplemented to prevent DNA damage (85). In a recent study, the addition of nucleosides have shown to have an important role in stem cell reprogramming by reducing the number of genomic rearrangements during the reprogramming of somatic cells (86). Although the importance of nucleosides have been demonstrated, the presence of nucleosides are typically at very low quantity as they also make up only 5-10% of the cell mass (87).

In most cell culture processes, amino acids are typically the most varied depending on the cell type and cell stage. Amino acids are categorized into essential and non-essential amino acids. In some cells, they lack the biosynthetic pathways to make certain amino acids and are specified as essential amino acids that must be supplemented to meet the cell's demand for it. This is key in medium design to understand the type of cell you have in culture and its biosynthetic capability. In certain tissues in-vivo, there are transport of

amino acids from one tissue to the next but in a more homogenous population of cells, they are limited by the biosynthetic pathways that are expressed in the cells. For instances, glutamine synthetase is typically expressed in cells in animals but after isolation this enzyme decreases after culture. As a result, the level of glutamine synthesis goes down requiring supplementation of glutamine into many cell culture media. Glutamine is an essential amino acid that must be provided as it is typically the highest consumed amino acids of all the 20 amino acids present in the cell culture media. In culture, they are known to consume roughly 1/10 of the consumption rate of glucose. Through carbon tracing experiments, glutamine has been found to contribute to the TCA cycle as well as through lactate production. Although glutamine is a critical amino acid for energy source and nucleoside synthesis as well, excess glutamine can lead to detrimental effects as it is known to degrade in aqueous solution and release ammonium which is toxic to cells. Thus, the balance between cellular demand and supply must be planned out carefully in order to reduce the potential fatal effects to cells.

Chapter 3: Materials and Methods

3.1 Human Embryonic Stem Cell Culture

The hES cell line H9 was predominantly used in all the experiments described in this thesis. However, it should be noted that there were other cell lines, HSF6 and HES3, that were also cultured to validate our findings but the results will not be present in this thesis. Therefore, all methods described was specifically carried out on the H9 cell line.

The human embryonic stem cell (hESC) line H9 was cultured using 80% Knockout Dulbecco's modified Eagle's medium (DMEM) (Gibco/BRL), supplemented with 20% Knockout Serum Replacement (Gibco), 2.0 mM glutamine, 0.1mM nonessential amino acids (Gibco), 55 mM β-mercaptoethanol, and basic fibroblast growth factor (FGF2; 10 ng/mL; R&D Systems). The cells were cultured on irradiated E13-E14 CF-1 mouse embryonic fibroblasts (MEFs; Charles River Laboratories, Wilmington,MA), at 37°C in 10% CO2. The cells were passaged on a regular basis every 2–3 days using 0.1% (w/v) Collagenase Type IV (Gibco) in Knockout DMEM after reaching ~50% to 70% confluency.

In the metabolic studies, hESCs were cultured in feeder-free conditions to minimize the effects of feeder cells. hESCs were harvested/collected using 0.5mM EDTA (Corning) and cultured on 2% Matrigel® (BD Biosciences) coated plates in TeSR™-E8™ media (STEMCELL-Technologies).

3.2 Hepatocyte Differentiation

H9 cells were plated in 12-well plates coated with 2% Matrigel_ (BD Biosciences) in mTeSR medium (STEMCELL Technologies) for 24 h or until confluency was reached. Differentiation was initiated by switching to differentiation medium consisting of a 60/40 (v/v) mixture of low glucose DMEM (Gibco) and MCDB-201 (Sigma), supplemented with 26 µg/mL ascorbic acid 3-phosphate (Sigma), linoleic acid and bovine serum albumin (LA-BSA; Sigma; 0.25µg/mL BSA and 2.35 µg/mL linoleic acid), insulin-transferrin-selenium (ITS; Sigma; 2.5 µg/mL insulin, 1.38 µg/mL transferrin, 1.25 ng/mL sodium selenite), 0.4 µg/mL dexamethasone (Sigma), 4.3 µg/mL b-mercaptoethanol (Hyclone), 100 IU/mL penicillin, and 100 mg/mL streptomycin (Gibco). Two percent fetal bovine serum was added to the media (v/v) in Stage I for the first 6 days and 0.5% (v/v) for the remaining period. The differentiation medium was supplemented with stage specific growth factors:

Stage 1, Activin A (100 ng/mL) and Wnt3a (50 ng/mL); Stage 2, FGF2 (10 ng/mL) and BMP4 (50 ng/mL); Stage 3, FGF8b (25 ng/mL), FGF1 (50 ng/mL), and FGF4 (10 ng/mL); and Stage 4, HGF (20 ng/mL) and Follistatin (100 ng/mL). Differentiations were carried out at 21% O2 and 5% CO2 with a 50% media change every 2 days during differentiation. A complete media change was performed when changing the stages of differentiation (ie, day 6, 10, and 14).

3.3 Endodermal Differentiation on Microcarriers

Cytodex 3, a crosslinked dextran bead with a thin layer of denatured collagen was used for providing a matrix for the growth and differentiation of adherent cells. Cytodex 3 (GE Healthcare) was purchased as dry powder and prepared by washing 2g of microcarrier powder with PBS for 3 times. Swollen beads were suspended in 500 ml PBS and sterilized by autoclaving. Different coating materials were used to coat Cytodex 3 and Matrigel™ was found to give the highest attachment efficiency for endodermal cells (Data not shown). Prior to differentiation, sterile microcarriers were coated with Matrigel™ for 1h at 37°C. Coated microcarriers were washed with low glucose DMEM (Gibco) twice and incubated with Stage 2 medium. Differentiated endodermal cells were washed with PBS and harvested using 0.1% (w/v) Collagenase Type IV (Gibco) in Knockout DMEM (Gibco) as described previously (78). Endodermal cells were suspended in Stage 2 medium and added to 12mg/ml of Cytodex 3 at the concentration of 1 x 10^6 cells/ml in a 1.7 ml tube (Corning). The mixture was incubated for 3 hours at 37°C to allow for initial cell attachment to the beads with occasional shaking of the tube every half an hour. The cell suspension was transferred to an ultra-low attachment plate (Corning) at a starting cell concentration of 1 x 10^6 cells/ml with 2 mg/ml Cytodex 3 microcarriers. After three days of differentiation and expansion in stage 2 medium, cells were further differentiated towards hepatocytes as described previously (raju et al. 2016)

3.4 Endodermal Cell Harvest and Expansion

For endodermal cell expansion, cells were harvested on Day 6 after the initiation of hepatocyte differentiation using 0.1% (w/v) Collagenase Type IV (Gibco) in Knockout DMEM and plated onto Matrigel pre-coated plates in Stage 2 medium containing FGF2 (10 ng/mL) and BMP4 (50 ng/mL). A day after plating, the media was completely

replenished with Stage 2 medium to remove any unattached cells. On the 3rd day after plating, the cells were passaged again using 0.1% (w/v) Collagenase Type IV onto Matrigel coated plates in Stage 2 medium. This was demonstrated to be capable of up to 3 passages without dedifferentiation of endodermal cells into fibroblastic cells. Three days after the second passaging, the medium was changed to Stage 3 medium to continue the differentiation.

3.4 Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)

Total RNA was obtained from cell lysates using the RNAeasy Micro Kit (Qiagen). Briefly, cells in culture are lysed directly on the tissue culture dish surface or in a 15 ml Falcon tube for cells not attached. The lysing procedure involves aspirating the media, rinsing the cells once in PBS, and finally lysing the cells using RLT lysis buffer + 1% βME, as prescribed by Qiagen. Each sample is collected in 350 µL of lysis buffer, vortexed for 30 second to homogenize the sample, and finally stored in a –80°C freezer for later processing.

Complementary DNA (cDNA) was synthesized using the Superscript III Reverse Transcriptase Kit (Invitrogen) according to manufacturer's instructions. Transcript abundance levels of a sample were normalized to the housekeeping gene, GAPDH. Gene expression is quantified relative to a housekeeping gene, typically GAPDH, and is typically expressed as either expression relative to GAPDH, $\log_2$(expression relative to GAPDH), fold change relative to a reference, or $\log_2$(expression relative to a reference). The following formulas are used to calculate these values:

$$\Delta Ct_{t,gene} = Ct_{t,gene} - Ct_{t,GAPDH}$$

$$\text{Fold change relative to GAPDH} = 2^{-(\Delta Ct_{t,gene})}$$

$$\Delta\Delta Ct_{t,gene} = \Delta Ct_{t,gene} - \Delta Ct_{t=ti,gene}$$

$$\text{Fold change relative to time } t_i = 2^{-(\Delta\Delta Ct_{t,gene})}$$

Sequences for the primers of the different genes used in this study are listed in Table 3-1. For comparison among cells at different stages of differentiation, some values were then transformed to be expressed as log10 (Expression level relative to hESC) instead of $\log_2$.

Table 3-1. Primer Sequence

Genes	Forward sequence	Reverse sequence
AAT (H9)	TTTAAAGGCAAATGGGAGAG	CCTAAACGCTTCATCATAGG
AAT (HSF6)	GTCAAGGACACCGAGGAAGA	TATTTCATCAGCAGCACCCA
AFP (H9)	CCTACAATTCTTCTTTGGGCT	AGTAACAGTTATGGCTTGGA
AFP (HSF6)	AAATGCGTTTCTCGTTGCTT	GCCACAGGCCAATAGTTTGT
ALB	TGGCACAATGAAGTGGGTAA	CTGAGCAAAGGCAATCAACA
CXCR4	AACTTCAGTTTGTTGGCTGC	GAAACAGGGTTCCTTCATGG
CYP3A4/5/7	AAGTCGCCTCGAAGATACACA	AAGGAGAGAACACTGCTCGTG
CYP3A7	TGCTTTGTCCTTCCGTAAGGG	CAGCATAGGCTGTTGACAGTC
CYP7A1	CTGAGGCTTTCCAGTGCCT	AGGTAGTCTTTGTCTTCCCGT
E-CADHERIN (HSF6)	CGAGAGCTACACGTTCACGG	GTGTCGAGGGAAAAATAGGCTG
E-CADHERIN (H9)	CGAACTATATTCTTCTGTGAGAGG	GATAGATTCTTGGGTTGGGTC
FOXA2	ATTGCTGGTCGTTTGTTGTG	TACGTGTTCATGCCGTTCAT
G6PC	GTGTCCGTGATCGCAGACC	GACGAGGTTGAGCCAGTCTC
GAPDH (H9)	TGGTATCGTGGAAGGACTCATGAC	ATGCCAGTGAGCTTCCCGTTCAGC
GAPDH (HSF6)	GAGTCAACGGATTTGGTCGT	GACAAGCTTCCCGTTCTCAG
GSC	TCTCAACCAGCTGCACTGTC	CCAGACCTCCACTTTCTCCTC
HAND1	CCATGCTCCACGAACCCTTC	CCTGGCGTCAGGACCATAG
HNF1α	GTGGCGAAGATGGTCAAGTCC	CCCTTGTTGAGGTGTTGGG
HNF1β	AGGCCACAATCTCCTCTCAC	TTGCTGGGGATTATGGTGGGA
HNF4α	TGTACTCCTGCAGATTTAGCC	CTGTCCTCATAGCTTGACCT
MIXL1	GGATCCAGGTATGGTTCCAG	CATGAGTCCAGCTTTGAACC
NKX6.1	TCAGGTCAAGGTCTGGTTCC	TCAACAGCTGCGTGATTTTC
OCT4 (H9)	GATGGCGTACTGTGGGCCC	TGGGACTCCTCCGGGTTTTG
OCT4 (HSF6)	CTTCGCAAGCCCTCATTTC	CCTTGGAAGCTTAGCCAGGT

SOX17	CGCACGGAATTTGAACAGTA	GGATCAGGGACCTGTCACAC
SNAI2	CGAACTGGACACACATACAGTG	CTGAGGATCTCTGGTTGTGGT
TTR (H9)	AAACCAGTGAGTCTGGAGAG	CTGTGAATACCACCTCTGCA
UGT1A1	CAACTGCCTTCACCAAAATCCA	GCAAGATTCGATGGTCGGGTT
MPC1	ACT ATG TCC GAA GCA AGG ATTTC	CGCCCACTGATAATCTCTGGAG
MPC2	TACCACCGGCTCCTCGATAAA	TATCAGCCAATC CAG CAC ACA
HK1	GCTCTCCGATGAAACTCTCATAG	GGACCTTACGAATGTTGGCAA
HK2	GAGCCACCACTCACCCTACT	CCAGGCATTCGGCAATGTG
HK3	GGACAGGAGCACCCTCATTTC	CCTCCGAATGGCATCTCTCAG
GCK	CCTGGGTGGCACTAACTTCAG	TAGTCGAAGAGCATCTCAGCA
GLS2	AACGAATCCCTATCCACAAGTTCA	GCAGTCCAGTGGCCTTTAGTG
GLS1	TGGTGGCCTCAGGTGAAAAT	CCAAGCTAGGTAACAGACCCTGTT T

3.5 Immunofluorescence

Cells were either grown on tissue culture plates or on glass coverslips that were placed inside tissue culture plates in the cases of higher magnification microscopes. The presence of proteins are visualized under a fluorescent microscope by marking the proteins of interest with fluorescent dyes attached to specific antibodies or secondary antibodies that interact with primary antibodies for the proteins. Cells in culture are washed once with PBS and then fixed for 15 minutes at room temperature with 4% formalin.

Fixed cells were washed with PBS once and then blocked with phosphate buffered saline (PBS) containing 0.2% Triton X-100, and 1% donkey serum or BSA (Sigma) at room temperature for 1 h. After blocking, samples were incubated with primary antibodies AFP (1:1,000; Dako), ALB (1:1,000; Dako), SOX17 (1:20; R&D Systems), Forkhead Box A2 (FOXA2; 1:1,000; Abcam), AAT (1:,1000, Abcam) and DAPI (1:500; Life Technologies) overnight at 4°C. The cells were then incubated with secondary antibodies [anti-mouse IgG1 Alexa Fluor 488 labeled (1:500 dilution;Molecular Probes), anti-rabbit IgG Alexa Fluor 488 (1:500 dilution; Molecular Probes), or anti-mouse IgG Alexa Fluor 555 (1:500 dilution; Molecular Probes)] for 30 min at room temperature. Negative controls

were cells incubated with only the relevant isotype control and secondary antibody. Dilutions and vendor of each antibody are provided in Table 3-2. Images were taken using an Olympus FluoView FV1000 BX2 confocal microscope at the University Imaging Center (Minneapolis, MN) or a Zeiss AxioVert 200M inverted fluorescence microscope at the Stem Cell Institute.

Table 3-2. Antibody Concentration and Vendor Information

Name	Company	Catalog no.	Dilution	Type
AFP	Dako	A0008	1:1000	Primary
ALB	Dako	A0001	1:1000	Primary
FOXA2	Abcam	ab40874	1:1000	Primary
SOX17	Abcam	ab84990	1:20	Primary
AAT	Abcam	Ab9373	1:1000	Primary
Goat IgG	Jackson Immunoresearch	005-000-003	Same as primary	Isotype
Rabbit IgG	Jackson Immunoresearch	011-000-003	Same as primary	Isotype
Donkey anti-goat Alexa 555 (red)	Invitrogen	A-21432	1:500	Secondary
Goat anti-rabbit Alexa 555 (red)	Invitrogen	A-21429	1:500	Secondary
Goat anti-rabbit Alexa 488 (green)	Invitrogen	A-11008	1:500	Secondary
Goat anti-mouse Alexa 488 (green)	Invitrogen	A-21121	1:500	Secondary
Hoechst 33258	Sigma	33258	1:500	Secondary

3.6 Mass Cytometry

Antibodies were purchased from R&D systems and conjugated with different elemental isotopes with different masses according to the manufacturer's protocol using the MaxPar® antiobidy labeling kit (Fluidigm). The concentration of the metal conjugated antibodies was measured using the absorbance reading at 280nm and diluted to 0.5 mg/mL in antibody stabilizer PBS (Candor Biosciences) with 0.05% sodium azide and stored in 4°C until further use.

Lyophilized antibodies for FOXA2, SOX17, AFP, ALB, DLK1, A1AT and HNF4α proteins were obtained from R&D systems. About 200 μg of antibodies were conjugated with a selected panel of heavy metal isotopes. The antibodies were eluted in 50 μl PBS instead of W buffer provided in the kit and this was found to improve antibody recovery.

Metal conjugated CXCR4 and CD44 were directly obtained from DVS Sciences. Cells were dissociated into single cells by treatment with 0.1% collagenase in DMEM and trypsin as described previously. About 500,000 cells per time point were fixed using 10% formalin for 20 minutes at room temperature. The cells were washed with PBS, centrifuged and each cell pellet were suspended in 5 µL Human TruStain FcX™ (BioLegend, 422302) and 95 µL PBS at room temperature for 10 minutes for blocking. The cells were first incubated in a cocktail of metal conjugated antibodies targeting surface markers in 100 µl of PBS for 30 minutes at room temperature. The concentration for each marker used is provided in Table 3-2.Cells were washed twice with PBS and incubated with a second cocktail of antibodies targeting intracellular proteins suspended in 100 µl SAP buffer (PBS with 0.1% (w/v) saponin (Sigma, 47036) and 0.05% (w/v) sodium azide (Sigma, 438456) for 30 minutes at room temperature. After washing the cells twice with SAP buffer, cells were incubated with MaxPar® Intercalator-Ir 125 µM (DVS Sciences, 201192A) at a dilution of 1:1000 in 1 ml of SAP buffer overnight at 4°C. Cells were washed twice, suspended in 500 µl water and passed through cell strainer and were run on the CyTOF2 instrument (DVS Sciences). During titrations, an antibody cocktail was made at 8 µg/ml and serially diluted twofold to achieve a wide range of concentrations. The negative control used was hESCs. It was critical to run a negative control for every titration as well as experiment in case of changes in antibody activity. Data was analyzed using the Cytobank software and visualized in Spotfire (Tibco).

Table 3-3: Antibodies for Mass Cytometry

Antigen	Vendor	Catalog number	Metal Conjugated	Concentration
SOX17	R&D	AF1924	Sm154	4 µg/ml
FOXA2	R&D	AF2400	Gd156	4 µg/ml
AFP	R&D	AF1369	Nd143	4 µg/ml
DLK1	R&D	MAB1144	Tm169	4 µg/ml
ALB	R&D	MAB1455	Er166	4 µg/ml
A1AT	R&D	AF1268	Tb159	4 µg/ml
CXCR4	DVS	3175001B	Lu175	2 µg/ml
CD44	DVS	3171003B	Yb171	2 µg/ml

3.7 Analysis of Mass Cytometry using Spanning-Tree Progression Analysis of Density Normalized Events (SPADE)

FCS files were obtained for each sample from the CyTOF2 and reanalyzed in the FlowJo software. For further analysis of mass cytometry and carry out clustering at the single cell level, each fcs file for each sample was concatenated and normalized by assigning the same DNA intercalator median intensity value for each sample. A metal-encoded DNA intercalator, Ir191 and Ir193, was used to label nucleated cells. Events with very low intercalator reading (<20) or very high (>5000) were removed to exclude cell debris and cell aggregates. The normalized fcs files were analyzed and visualized further using the Cytobank software (Fluidigm). For Spade analysis, arcsinh was used to transform the intensity and cluster into 75 nodes. Because SPADE analysis is clustering based on the value of intensity, the SPADE tree can be annotated in a quantitative manner by labeling high, moderate, and low expression for certain markers. The SPADE trees were then overlaid for each sample to show the population distribution for that particular sample.

3.8 Functional Analysis of HLCs

Several functional assays were carried out to determine the functional maturity of HLCs. Albumin secretion was measured using a quantitative ELISA kit (Bethyl Laboratory) following the manufacturer's instructions. An internal standard, provided with the kit, was used to obtain a four-parametric equation to enable the best estimation of the albumin concentration in our samples. The albumin concentration in the sample was then subtracted from the amount present in the fresh media to quantify the amount secreted by the cells after 1 or 2 days of incubation. The cell number was measured by trypan blue exclusion and the albumin secretion was reported as pg/cell/day.

Urea secretion was also measured by using an QuantiChrom Urea Assay Kit (BioAssay Systems) according to the manufacturer's instruction. Briefly, cells were rinsed with PBS and then grown in the hepatocyte differentiation media with 1mM ammonia-bicarbonate with 0.5% FBS for 24 h. Supernatant was collected and measured using the kit and reported on a per cell basis as well.

Periodic Acid-Schiff staining was performed using a kit from Sigma Aldrich by immersing the cells in the Periodic acid solution for 5 mins and then in the Schiff's reagent

for 15 minutes. The cells were then washed with PBS before taking it to microscopy for visualization.

To test for the function of drug metabolism, several CYP450 enzymes were measured at the transcript level after induction through the exposure of rifampicin. Cells were washed with PBS and then immersed in media with 50 µM rifampicin (Gold Biotechnology. The transcript levels of *CYP2A6, CYP2C8,* and *CYP2C9* were measured using qRT-PCR.

3.9 Transcriptome Analysis

Total RNA was extracted from different samples during hESC differentiation at different time points of differentiation as well as endodermal cell expansion using the RNease Mini kit. Samples were hybridized to the Illumina HT12 bead array v3 (Illumina Inc). CEL files were processed using RMA summarization and transcriptome data from over 34,000 probes representing about 20,000 genes were obtained. Data was processed using the *lumi* pacakage in R. First, intensity values were condensed to a single value based on common ENSEMBL ID. Next, data was normalized using either linear or quantile normalization in MatLAB where the linear normalization was set to have 500 as the median intensity. Briefly, principal component analysis (PCA), hierarchical clustering, and NMF was performed in R. The algorithm and criteria will be described more in depth in later sections.

3.10 Western Blot

Cells were washed with cold PBS before being harvested and lysed using cold RIPA buffer (Abcam) with a cold plastic cell scraper. Cells were stored on ice in RIPA buffer and centrifuged in a microcentrifuge at 4°C for 30 min at 12,000 rpm. The supernatant was transferred to a fresh new tube and stored on ice, where the cell pellet was discarded. Protein quantification was carried out using the Pierce™ BCA Protein assay kit (Thermo #23227) following the manufacturer's instructions. To reduce and denature the sample, NuPage™ LDS Sample buffer (4X, Thermo) was added depending on how much protein to load along with 1ul of β-mercaptoethanol. Each cell lysate was boiled in the buffer at 100°C for 5 min. 15 ug of protein were separated using Bolt™ 4-12% Bis-Tris Plus gels (Thermo). The primary antibody used were anti-COX1(1:5000, Cell Signaling

Technology), anti-COX2(1:5000, Cell Signaling Technology),anti-COX3 (1:2000, Abcam) , and anti-βActin (1:1000,Cell Signaling Technology). The secondary antibody used was goat anti-rabbit IgG HRP (1:50,000, Santa Cruz). For signal development, the SuperSignal West Pico Chemiluminescent Substrate kit (Thermo Scientific) was used.

3.11 Mitochondria Biogenesis

Mitochondria DNA was measured using a Mitochondria DNA isolation kit (AbNova KA0895). Briefly, cells were harvested using Collagenase Type IV and centrifuged at 600g for 5 minutes at 4°C. Cells were resuspended in 1X Cytsol Extraction Buffer and incubated on ice for 10 min. Cells were homogenized using a tissue grinder. The homogenate was centrifuge at 600g at 10 min at 4°C. The supernatant was transferred to a new tube and centrifuged at 10,000g for 30 min at 4°C. The pellet was resuspended in 1ml 1X Cytsol Extraction Buffer and centrifuged at 10,000g for 30 mins at 4°C. The supernatant is removed and the pellet is lysed with mitochondria lysis buffer and kept on ice. Enzyme B was added and incubated at 50°C until solution becomes clear and then absolute ethanol is added and stored in -20°C for 10 min. The solution is centrifuged and the pellet can be resuspended in water for quantification using a nanodrop at 250nm.

3.12 Metabolite Measurement

Supernatant was collected every other day for assessing the concentration of different metabolites.

3.12.1 Glucose Measurement

Glucose was measured using Infinity Glucose Hexokinase (Thermo) following the manufacturer's instruction. Briefly, for glucose concentration measurements, 196 µL of the hexokinase reagent solution (containing NAD, ATP, HK, G-6-PDH) was pipetted into a 96 well plate. 4 µL of distilled water (for blank), standard solution (1-0.0125 g/L glucose), or sample was then added, the plate was shaked to allow the mixture to mix. The reaction was allowed to be completed for 5 minutes at room temperature and the absorbance was measured at 340 nm against distilled water as reference. The glucose concentration of the unknown sample was then given by:

$$\frac{A_{sample} - A_{blank}}{A_{std} - A_{blank}} \times dilution \times Std.Conc$$

where

$$\Delta A = A_{initial} - A_{final}$$

The control solution contained 1 g/L substrate and was serially diluted to create a standard curve.

3.12.2 Lactate Measurement

Lactate measurements were carried out using a lactate assay kit (Sigma) following the manufacturer's instruction. Briefly, a lactate standard was created using the 100nmol/µl lactate standard and diluted with lactate assay buffer to create a 1nmol/µl solution. The reaction mix containing lactate assay buffer, lactate enzyme mix, and lactate probe was mixed with the sample and allowed to incubate in the dark at room temperature for 30 min. The reading was measured at 570nm and the concentration was determined based on the equation from the standard curve.

3.12.3 Amino Acid Measurement

Amino acid concentration was measured by HPLC (Agilent) using a 4.6 x 150mm Zorbax Eclipse Plus C-18 column (Agilent) with the derivatization reagents, OPA(Agilent) and FMOC (Agilent). Cysteine, tryptophan and proline cannot be measured by the OPA method and thus FMOC was purchased to resolved the remaining amino acids. The mobile phases consist of:

Mobile Phase A:

10mM Na_2HPO_4 (pH 8.2)

10mM $Na_2B_4O_7$

5mM NaN_3

Mobile Phase B:

Acetonitrile	45% v/v
methanol	45% v/v
water	10% v/v

The flow rate of the mobile phase was maintained at 1.0 mL/min and the gradient profile was as follows:

Table 3-4. HPLC Gradient Profile.

Time (minutes)	% A	% B
0 – 0.5	98	2
0.5	98	2
20	43	57
20.1	0	100
23.5	0	100
23.6	98	2
25	End	end

The exact concentration of the amino acids was determined by comparing the observed peaks with those generated from fresh media and its known amino acid concentration.

3.12.4 Ammonia Measurement

Ammonia was measured using an ammonia diagnostic kit (Sigma, St. Louis, MO for the determination of ammonia concentration for each sample. The assay is based on reductive amination of 2-oxoglutarate, using glutamate dehydrogenase (GLDH) and nicotinamide adenine dinucleotide phosphate (NADP) as follows:

$$\text{2-oxoglutarate} + NH_3 + \text{NADPH} \xrightarrow{\text{GLDH}} \text{glutamate} + \text{NADP}$$

The decrease in absorbance at 340 nm due to disappearance of NADPH is proportional to the ammonia concentration.

150 µL of ammonia assay solution (containing 2-oxoglutarate and NADPH) was pipetted into several wells of a 96 well plate. 10 µL of distilled water (for blank), control solution (for standard), or sample was added. The cuvettes were allowed to equilibrate for 5 minutes at room temperature. The initial absorbance of each cuvette at 340 nm was measured against distilled water as reference. 5 µL of GLDH solution was then added to each well and the reaction was allowed to occur for 5 minutes after gentle shaking. The final absorbance was again measured at 340 nm against distilled water as reference. The ammonia concentration was then given by:

$$\frac{\Delta A_{test} - \Delta A_{blank}}{\Delta A_{std} - \Delta A_{blank}} \times dilution \times Std.Conc.$$

where

$$\Delta A = A_{initial} - A_{final}$$

The control solution used had a concentration of 5 mg/L and a standard curve was generated through serial dilution.

3.12.5 Oxygen Consumption Measurement

Oxygen consumption was measured with the NeoFox Software viewer using the NeoFox-GT fluorimeter coupled with the FOSPOR-R Oxygen Sensor probe (NeoFox). Cells were harvested and submerged in an Eppendorf tube with no air bubbles to allow for oxygen diffusion. The change in oxygen level would be due purely to oxygen consumption by the cells.

3.12.6 Specific Consumption/Production Rates

Metabolic consumption/production rates were calculated based on the cell count for each day. The viable and total cell concentration was measured by trypan blue dye exclusion using a hemacytometer. 1ml of the sample was vortex to ensure homogenous mixing and then 10µL of the sample was diluted with 10 µL of the 0.1% trypan blue solution. The mixture was transferred to the hemacytometer and viewed under the microscope. The cell count is given by:

(number of cells x dilution) / (number of 1 mm^2 box) x 10^4 (cells/ml)

The specific consumption/production rates were determined by:

Consumption rate = $\frac{S_t V_t - S_{to} V_{to}}{\Delta t * Cell\ number}$ where S is the substrate concentration and V is the volume of media

Chapter 4: Endodermal Cell Expansion During Hepatocyte Differentiation

4.1 Introduction

Pluripotent stem cells possess the ability to provide a renewable source of hepatocytes upon differentiation overcoming the issues of donor shortages. In our laboratory, we had established a protocol capable of differentiating pluripotent stem cells to the hepatic lineage using a sequential treatment of growth factors and cytokines that mimic embryonic liver development. In the past few years stem cell research has made significant advances; stem cells and progenitors cells isolated from various sources can now be expanded in cell number and differentiated towards the liver lineage in vitro. This has brightened the prospect of generating large numbers of functional hepatocytes for applications in hepatic cell transplantation, extracorporeal liver-assist devices and in liver tissue engineering. Because of the prevalence of liver diseases, the severity of liver failure and its lack of curative treatments, it is imperative that we further advance stem cell based regenerative medicine for liver disease treatment. In addition to liver failure treatment, many congenital liver disease will also benefit from stem cell based therapy.

Our understanding of liver development has aided in different groups in developing different strategies to obtaining HLCs from stem cells (88-90). The differentiated cells by the end of every differentiation process result in cells with remarkably similar properties as primary hepatocytes. However, with the promises of stem cell in the field of regenerative medicine especially for chronic liver failure or liver diseases, low cell yield from a current differentiation is still a major hurdle that needs to be addressed. It is anticipated that any therapeutic applications requiring hepatocytes would be in the order of 10^9–10^{10} cells/treatment. To derive HLCs currently, stem cells are expanded to the required quantity before undergoing differentiation toward the hepatocyte lineage. An alternative strategy to circumvent this issue would be to not only expand stem cells in their pluripotent state but also devise a methodology to expand PSC-derived progeny during their differentiation toward hepatocytes. We hypothesized that with the appropriate signaling cues, it may be

possible to expand the number of cells during the process of in vitro hepatocyte differentiation by mimicking the native proliferation that occurs naturally in vivo.

To date, there has not been a report of simultaneous in vitro proliferation and differentiation of PSC-progeny, just like in vivo fetal liver development. There have been multiple reports on selective surface markers that could be used for isolation of renewable hepatocyte progenitor cells (91-93). In one study, hepatocyte progenitor cells were enriched using the surface markers EpCAM or N-cadherin and, subsequently, expanded on stromal feeder layers (93). In addition, a self-renewing endodermal cell line was reported where $CXCR4^{+}/CD117^{+}$ cells were sorted and expanded on mouse embryonic feeders (48). During liver development, cells can undergo extensive proliferation using the signaling cues from their surrounding environment to increase the liver mass. In this chapter, we demonstrated the expansion of an early stage of endodermal intermediates without the need for cell sorting or relying on feeder cells that were previously described in other studies. This method was still able to show the ability to achieve higher yields of HLCs after subsequent differentiation of these endodermal intermediates from the same starting cell number of hESCs.

4.2 Results

4.2.1 Expansion of Endodermal Cells

In this study, we showed that during hepatocyte differentiation from hESCs, an intermediate population could be expanded similar to in-vivo development and simultaneously differentiated to the hepatic lineage to give larger quantities of HLCs as shown in Figure 4-1. hEScs were differentiated to the definitive endoderm stage using a medium containing Activin and Wnt3a to reach cell densities of 2.5×10^5 cells/cm^2 in six days. The resulting endoderm cells were then detached by 0.1% collagenase treatment and passaged at ~6×10^4 cells/cm^2 onto Matrigel coated plates in Stage 2 medium containing FGF2 (10ng/ml) and BMP4 (50ng/ml). The morphology of the cells immediately after plating and before passaging are shown in Figure 4-2. Cells adhered to the surface a few hours after plating and expanded up to 4-fold in viable cell number after 3 days. Cells were then passaged again in Stage 2 medium containing FGF2 and BMP4, which have been

reported to provide the necessary proliferative cues to endodermal cells during embryonic liver development (94).

The endodermal cell population expanded approximately eightfold after two passages as shown in Figure 4-3. These populations will be referred to as EN1 and EN2 where the number denotes the passage number that these cells have undergone. Further passages beyond the second passage were carried out, resulting in cell expansion up to 15-fold; however, we detected an increasing population of cells with a fibroblastic morphology that would overtake the native population. By contrast, when we tracked the cell expansion during Stage 2 of the conventional differentiation method without passaging, we observed that the cell expansion was limited only up to twofold as shown in Figure 4-3. Thus, by implementing two passaging steps during the hepatic endoderm commitment stage, we could induce an eightfold expansion by providing additional surface area with the signaling cues of Stage 2 medium.

4.2.2 Expression of Hepatic Genes and Protein Levels During Expansion Process

We evaluated the expression of pluripotency, endoderm, and hepatic endoderm related genes in cells during the expansion by qRT-PCR and immunostaining. Expression of Octamer-binding transcription factor 4 (OCT4), a master regulator of the pluripotency network in hESCs (95), decreased about 1,000-fold during the course of endodermal expansion as shown in Figure 4-4. The endodermal transcription factor, goosecoid homeobox (GSC), a key marker for differentiating definitive and visceral endoderm (96), and CXCR4, a surface marker coexpressed with GSC (96), were both highly expressed in the D6 population, but decreased in the EN1 and EN2 populations (Figure 4-4). Our hypothesis was that similar to in vivo development, ESC-derived DE cells can proliferate while at the same time differentiate to hepatic endoderm. In addition, we saw a higher expression of endodermal markers and lower expression of pluripotent markers in the new expansion protocol possibly due to the preferential attachment of endodermal fated cells compared to undifferentiated hESCs.

We also evaluated the expression of the genes indicative of maturation to hepatic committed endoderm and hepatoblasts. In line with our hypothesis, the decrease in DE marker gene transcripts was accompanied by almost 1,000- fold increase of the transcript

level of alpha-fetoprotein (AFP) and albumin (ALB) over the expansion period (97). These dynamics of gene expression suggest the gradual transition of an endodermal-committed phenotype toward a more hepatic phenotype during the expansion process.

We also compared the transcript levels of hepatic endoderm/ hepatoblast genes in the EN1 and EN2 populations to cells obtained at the end of Stage 2 (D10) and Stage 3 (D14) from the conventional differentiation process without cell expansion. Expression levels of OCT4, CXCR4, and GSC in EN1 and EN2 cells were similar to those in D10 and D14 cells, respectively (Figure 4-4). Levels of the hepatic transcripts, AFP and ALB, increased more gradually in the conventional method of differentiation (D10 and D14), compared to cells undergoing cell expansion (Figure 4-4).

The transition from an endoderm to a more hepatic stage was also examined by immunostaining in both the conventional and expansion method as shown in Figure 4-5. FOXA2 and SRY (Sex determining Region Y)-Box 17 (SOX17), key transcription factors in the establishment of DE, were prominent in the D6 stage, but diminished by the EN2 and D14 stage. In contrast, AFP was absent in the D6 stage and became more prominent in both the EN2 and D14 population consistent with the transcript levels observed. The results suggest that cells undergoing expansion are differentiating simultaneously.

4.2.3 Mass Cytometry and SPADE Analysis

To identify whether the increase in transcript for hepatic endoderm/hepatoblast marker genes and proteins during expansion was restricted to a differentiated subpopulation of cells, or was occurring in most of the population, we used mass cytometry to examine coexpression of a panel of endodermal and hepatic markers at a single cell resolution (98). Cells undergoing differentiation were labeled with stable isotopes of lanthanide heavy metal conjugated with antibodies against endodermal marker proteins (CXCR4, FOXA2, and SOX17) and hepatic marker proteins (DLK1, CD44, AFP, ALB, and AAT).

Similar to flow cytometry, the antibody-labelled cells were dissociated into single cells and analyzed at a single cell resolution for multiple markers at a time. However, instead of detecting using various fluorescent tags, the sorted cells are vaporized to leave the stable isotope tags to be analyzed by a time of flight (TOF) mass spectrometry. In the

TOF analysis, different antibody tags will give sharp and distinctive signatures corresponding to its conjugated metal isotope, allowing for quantification of each labelled antibody without the need of resolving spectrum spillover or overcoming auto-fluorescence that normally occurs in multiple parametric flow cytometry.

Due to the complexity of the many markers that was used to label a particular cell, three-dimensional (3D) diagrams were used to efficiently plot the different marker expression levels. During the course of differentiation, various combinations of marker expressions were plotted as 3D diagrams with three markers represented along the X, Y, and Z axes and CD44 represented as a color gradient as shown in Figure 4-6 and 4-7. Still, with such plots it is not easy to gain a global view of the evolution of the population during the directed differentiation. Thus, we used the expression level of all eight markers to characterize the co-expression in the different populations using SPADE, which groups the cells into a defined number of clusters based on its expression pattern (99).

The SPADE analysis utilizes a combined dataset that consists of all the markers' expression level for cells of different differentiation stage during the directed differentiation as well as the expansion stages. The dataset was then clustered into 75 nodes based on the expression level of the eight markers for each cell, and the clustering results are shown as a tree diagram (Figure 4-8). Each node constitutes a cluster of cells with a similar pattern of expression for all eight markers. Because eight markers are used, eight tree diagrams are used to present the results for each marker (Figure 4-8).

Based on the expression pattern of the different lineage specific markers, certain nodes can then be assigned to one of three subpopulations: (A) endodermal cells, (B) hepatic endoderm/hepatoblasts, and (C) hepatoblasts/hepatocytes (Figure 4-8(a-h)). The designation of the different subpopulation is based on the expression pattern of the different markers. The following expression patterns were used to classify the subpopulations: (A) $(CXCR4^{++/+} \cup FOXA2^{++/+} \cup SOX17^{++/+}) \cap DLK1^{-} \cap AFP^{-} \cap ALB1^{-} \cap AAT^{-}$ ("++/+" denotes ranging from ++ to +, and "∪" and "∩" represent union and intersection, respectively). (B) encompasses three subclasses: (i) $(CXCR4^{++/+} \cup FOXA2^{++/+} \cup SOX17^{++/+}) \cap (DLK1^{++/+} \cup AFP^{++/+} \cup ALB^{++/+} \cup AAT^{++/+})$, (ii) $CXCR4^{++/+/-} \cap FOXA2^{++/+/-} \cap SOX17^{++/+/-} \cap (DLK1^{++/+} \cup AFP^{++/+}) \cap ALB^{+/-} \cap AAT^{+/-}$, or (iii)

$CXCR4^{++/+/-} \cap FOXA2^{++/+/-} \cap SOX17^{++/+/-} \cap DLK1^{++/+/-} \cap AFP^{++/+/-} \cap (ALB^{+} \cup AAT^{+})$. (C) $CXCR4^{-} \cap FOXA2^{-} \cap SOX17^{-} \cap DLK1^{++/+/-} \cap AFP^{++/+/-} \cap ALB^{++} \cup AAT^{++}$. The threshold for determining -/+/++ are described more in detail in the figure legend of Figure 4-8. The nodes of the same subpopulation are encircled in a dashed line. However, not all nodes were assigned to a subpopulation, especially those with low expression levels of most markers which we believe to be non-hepatic or endodermal committed cells.

To validate our classification of the nodes into the different subpopulations, we performed k-means clustering on the 75 nodes using the median intensity of each marker for each node. The clear majority of the nodes for subpopulations A (endodermal cells) and C (hepatoblast/hepatocytes) were assigned to its own distinct clusters. The majority of nodes in subpopulation B (hepatic endoderm/hepatoblasts) were mainly scattered across the two clusters (Figure 4-8i). The results of the k-means clustering illustrated that the majority of the nodes within a subpopulation were grouped to the same clusters supporting our classification of the 75 nodes into 3 subpopulations.

The mass cytometry data for each stage of differentiation was then mapped to the previously made SPADE tree diagram to illustrate the heterogeneity of the different subpopulations within each stage. The fraction of cells that were classified to a node for a particular stage will then be illustrated by the size and color of the node (Figure 4-9). Using the fraction of cells for each subpopulation, we can then quantitatively determine the heterogeneity at each stage of the differentiation/expansion. On D6, 51% of the cells were definitive endodermal cells, expressing CXCR4, SOX17, or FOXA2. During the expansion stage (EN1 and EN2), the definitive endodermal cells decreased to less than 10% while the early hepatic endoderm/hepatoblasts increased steadily to over 50%. The decrease of definitive endodermal cells and increase of hepatic endoderm/hepatoblasts was similar to what we observed in cultures without expansion (from D6 to D14). At the end of the directed differentiation, a similar percentage of HLCs was obtained with cell expansion (75%) or without (76%). However, considering that the cell number had increased significantly at the EN2 stage compared to D14 of differentiation, the total number of hepatocytes obtained is substantially higher with cell expansion.

We also used immunostaining to examine key differentiation markers (costaining of AFP and ALB or FOXA2 and SOX17) to confirm the results of mass cytometry measurements. The percentage of cells that were positive for those markers were compared to that obtained using mass cytometry for different populations (Figure 4-10). Images of multiple microscopic fields for each population sample were randomly taken, and the number of all nucleated cells and fluorescently stained cells were quantified using CellProfiler (100). Some differences were observed between the percentage of cells stained and percentage through mass cytometry. However, these small differences in percentages between the two approaches were likely caused by the chosen threshold values of mass cytometry data classifying positive and negative cells. Nevertheless, the overall trend of the dynamics of marker expression was still the same in both immunostaining and mass cytometry results.

4.2.4 Differentiation Potential and Functional Activity of Expanded Endodermal Cells

The results of the mass cytometry studies suggested that the majority of the expanded cells (EN2) can differentiate to HLCs. To confirm that HLCs derived with or without cell expansion were of the hepatic phenotype, both populations were immunostained and shown to be positive for the hepatic marker, ALB (Figure 4-11). We evaluated he transcript level of key hepatocyte marker genes in D20-HLCs and in EN2-HLCs, including phosphoenolpyruvate carboxykinase (PEPCK), alpha-1 antitrypsin (AAT), AFP, and ALB. The EN2-HLC population expressed all the hepatocyte transcript levels 10^2–10^5 higher compared to hESCs, similar to D20-HLCs (Figure 4-12a).

Further characterization of the HLCs and EN2-HLCs were carried out by measuring the activity of drug metabolism. Drug exposure to hepatocytes often induces specific CYP450 enzymes involved in its metabolism. Following exposure to rifampicin, we could detect increased transcript levels of its corresponding CYP450 enzymes (CYP2C8, CYP2C9, and CYP2A6) in D20-HLCs and EN2-HLCs (Figure 4-12b) (101). The level of increase was comparable between EN2-HLCs and D20-HLCs. In addition to drug metabolism, albumin synthesis and urea secretion for both D20-HLC and EN2-HLC

populations were also measured and shown to be comparable (Figure 4-13a-b) (102). Glycogen synthesis was evident in both D20-HLC and EN2-HLC cells as assessed by Periodic acid-Schiff staining (Figure 4-13c).

4.2.5 Comparative Transcriptome Analysis of Expansion and Differentiation

The transcriptome of cells at different stages of directed differentiation was evaluated and subjected to PCA. The first two principal components (PCs) representing 90% of the variance among all samples were chosen to plot all the samples in a two-dimensional PC space as shown in Figure 4-14. ESC, D6, D10, D14, and D20-HLC lined up chronically in order according to their differentiation stages, with the least and most mature cells on the two separate ends of the plot. EN1, EN2, and D10 cells, all exposed to Stage 2 medium, cluster in a similar region, with EN1 and EN2 lying between D6 and D10 cells. Both EN2-HLC and D20-HLC colocalize in the same region, almost overlapping one another, confirming their similarity despite the differences with respect to cell expansion. The gene expression dynamics of cells undergoing directed differentiation from both protocols was compared and plotted in groups of similar functional classes. Overall, the gene expression profiles show that, regardless of whether cells were expanded during the endodermal stage, their gene expression dynamics was very similar. The transcript level of pluripotent genes POU5F1, NANOG, SOX2, ZIC3, SALL4, and LIN28 decreased from the beginning of differentiation as expected (Figure 4-15). Endodermal genes EOMES, CXCR4, GSC, and SOX17 increased in the first stage of differentiation, then decreased in the following stages either accompanied by cell expansion or not (Figure 4-16).

In the presence of Stage 2 medium, a gradual increase in the transcripts of hepatoblast and hepatic genes is seen during the course of differentiation, including transcripts for AFP, DLK1, TTR, and KRT8 (hepatoblast specific markers), as well as ALB, PEPCK, PCK2, transferrin (TF), and a1-antitrypsin (A1AT; more mature hepatocyte markers) (Figure 4-17). The expression level of genes pertaining to important liver functions such as glycolysis/gluconeogenesis, urea cycle, fatty acid metabolism, and drug detoxification was also similar in D20-HLCs and EN2-HLCs (Figure 4-18). The similarity in transcript levels of these genes gives further credence to the notion that although the

endodermal cells underwent expansion, they retained undiminished differentiation capability to HLCs to a degree similar to that of unexpanded endodermal cells.

4.3 Discussion

The differentiation of PSCs toward the hepatocyte lineage raises the possibility of having an unlimited supply of human hepatocytes for a number of cellular applications. Although all protocols vary somewhat, they typically consist of treating stem cells with a combination of growth factors/cytokines in stages to mimic the key signaling cues occurring during embryonic development. However, cellular proliferation, which is integral during in vivo development, is typically absent or minimal in directed differentiations of stem cells to hepatocytes. In this study, we purposefully introduced cell expansion during the differentiation process between the time of commitment to DE and the time hepatic endoderm/hepatoblasts are being generated, with a goal to not only generate larger quantities of HLCs but also to more accurately mimic in vivo development where expansion also occurs.

In this study, we demonstrated that hESCs, during the course of hepatocyte differentiation, could be expanded at an intermediate endodermal stage and continued to differentiate toward a hepatocyte-like phenotype. To mimic the proliferative capability of developing hepatic cells, we utilized the soluble factors known to induce endoderm differentiation and promote proliferation as a way to incorporate an expansion stage. The factors, BMP4 and FGF2, were chosen to mimic the signaling of the cardiogenic mesoderm and septum transversum mesenchyme during the transition of the foregut endoderm to the liver bud, a process that involves extensive cell migration, proliferation, and differentiation.

By providing more surface area to cells at the intermediate endodermal stage, cells demonstrated the ability to expand up to at least eightfold. Given that we split the cells at a ratio of 1:4 during the first passage, the initial EN population showed remarkable proliferative capability within 3 days. By the third passage, up to 15-fold increase in cell number was occasionally achieved, but the resulting expansion from the third passage of cell expansion was not as consistent. In addition, cells with a fibroblastic morphology

began to emerge when cells were passaged beyond the second expansion stage, suggesting that the expanding endodermal cells have a limited proliferative capability similar to developing cells in vivo.

However, the increase in cell number during the expansion demonstrated that stem cell–derived endodermal cells exhibit a proliferative capability also seen with endodermal cells in developing embryos (103-105). The expansion in cell number was also accompanied by a simultaneous increase in hepatocyte transcript levels along with a drastic decrease in endodermal and pluripotent marker gene expression. This finding correlates with what happens to developing endodermal cells undergoing proliferation as they commit to the different endoderm lineages during liver development (103).

Our results indicated that the DE cells were capable of simultaneous proliferation and differentiation. However, it is not clear whether the changes in gene expression happen to a larger population of cells or to only a small population. We used mass cytometry to examine the progression of cells at different stages of differentiation. The population of cells at different stages was further classified into different subpopulations to examine the dynamics of the cells during differentiation. In the D6, EN1, EN2, and D14 populations, only 50%–75% of the population were categorized as endodermal cells, hepatic endoderm/hepatoblasts, or hepatoblasts/hepatocytes. However, by the end of the differentiation, the percentage of hepatic-committed cells had increased to almost >90%.

These results suggest that cells differentiate at different rates, giving rise to a heterogeneous population at different stages of differentiation. It is noted that a large number of the nodes in the SPADE tree were not included as the different classes of differentiated cells. Although the density of cells in each of those nodes was small, altogether they constitute a significant fraction of the total cells. It is likely that those cells are in a transition stage and also possible that they might include some cells that failed to differentiate toward the hepatic lineage. Nevertheless, the majority of the population is differentiating along the hepatic lineage. The decrease in endodermal and increase in hepatocyte markers occurred in a majority of the cells, as a large fraction of the population

coexpressed multiple hepatocyte markers and lost the expression of early stage markers simultaneously.

Furthermore, the results demonstrate that the increased hepatic transcript level in EN2-HLC was not merely caused by a small subset of cells expressing extremely high levels of hepatocyte markers, but was attributed by a majority (>75%) of the cells differentiating toward the hepatic lineage. Moreover, these novel quantitative analyses of the differentiation of endodermal cells confirmed that the majority of cells following the expansion process were also capable of becoming HLCs. Indeed, the resulting EN2-HLC population was similar to D20-HLC: both containing a large fraction of cells expressing moderate to high levels of ALB, AFP, AAT, and CD44, but low levels of endodermal markers. The similarity between EN2-HLC and D20-HLC was further confirmed by their ability to secrete albumin, synthesize urea, and induce cytochrome P450 enzymes in response to rifampicin.

Mass cytometry have been used to explore the heterogeneity in the hematopoietic system and in cancer (98, 99, 106-110). In stem cell research, it has been used to explore the emergence of induced PSCs from fibroblasts (111). In this study, we used mass cytometry to examine the population dynamics during directed differentiation of stem cells toward the hepatic lineage. Importantly, we took advantage of the detailed analysis at the single cell level provided by mass cytometry by imposing quantitative criteria for each marker to perform quantitative analysis of the emergence of different subpopulations at different stages of differentiation. The mass cytometry data demonstrate that while endodermal cells underwent expansion, they continued on the trajectory of differentiation toward the hepatocyte lineage. This was further supported through microarray analyses. Using PCA on the transcriptome data, we plotted the results on a PC1 versus PC2 space to demonstrate the similarities between each sample. All cell samples lined up in the order of their progression toward the hepatocyte lineage. EN1 and EN2 were positioned in between D6 and D10 cells, suggesting that their differentiation status is more advanced than D6, but not as mature as D10. Thus, during expansion, endodermal cells were proliferating while also differentiating toward a more hepatic endoderm/hepatoblast fate instead of remaining in a renewable endoderm cell stage (48).

In this study, hESCs were simultaneously expanded while differentiating toward hepatic lineage. Other studies have reported the expansion of hepatocyte progenitor cells. By FACS sorting, EpCAM+/c-KIT- hepatic progenitor cells from differentiating mouse ESCs were isolated and, subsequently, expanded on mouse embryonic feeders (92). In another study, human hepatoblasts were sorted from ESC progeny based on N-cadherin expression and, subsequently, cultured as progenitors on stromal feeder cells (48). Human endodermal progenitor cells have also been clonally isolated from a differentiating population based on expression of the surface markers CXCR4 and CD117 and expanded on Matrigel and feeder cells. These progenitor cells were subsequently differentiated to express hepatic markers. In all three studies, the cells were maintained at a ‘‘progenitor’’ state and did not appear to continue to differentiate until they were removed from the feeder and exposed to hepatocyte differentiation conditions. By contrast, in our study, the bulk of the starting population was expanded and differentiated simultaneously, mimicking embryonic liver development in which hepatic endoderm and early hepatoblast differentiation occurs while the size of the liver is rapidly increasing.

In spite of the increased expression of a number of hepatic genes, both D20-HLCs and EN2-HLCs are not yet mature hepatocytes as transcript levels of enzymes involved in xenobiotic biotransformation are significantly lower in HLCs than in neonatal or adult hepatocytes (Data not shown). In addition, the degree of cell expansion achieved in this study is still limited. For potential applications in the future, a substantially larger level of expansion will still be more desirable. Further improvement of the protocol for hepatocyte differentiation from PSCs is necessary to extend cell expansion and to achieve maturity of HLCs. However, our findings indicate that it is possible to expand endodermal cells during hepatocyte differentiation, which might alleviate the time and effort in generating large quantities of HLCs.

4.4 Figures

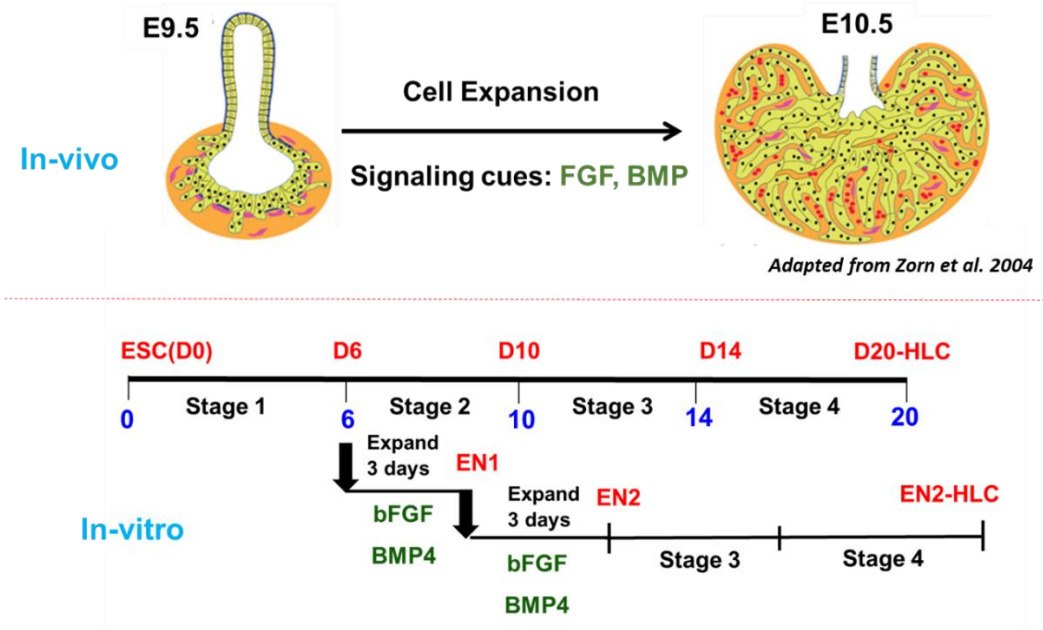


Figure 4-1. Protocol of Expansion of Endodermal Cells During Hepatocyte Differentiation. A) Cells were differentiated towards the endoderm stage and subsequently passaged using a combination of bFGF and BMP4 similar to developmental cues during cell expansion. The expanded cells were then further differentiated towards the hepatic lineage using the existing differentiation protocol.

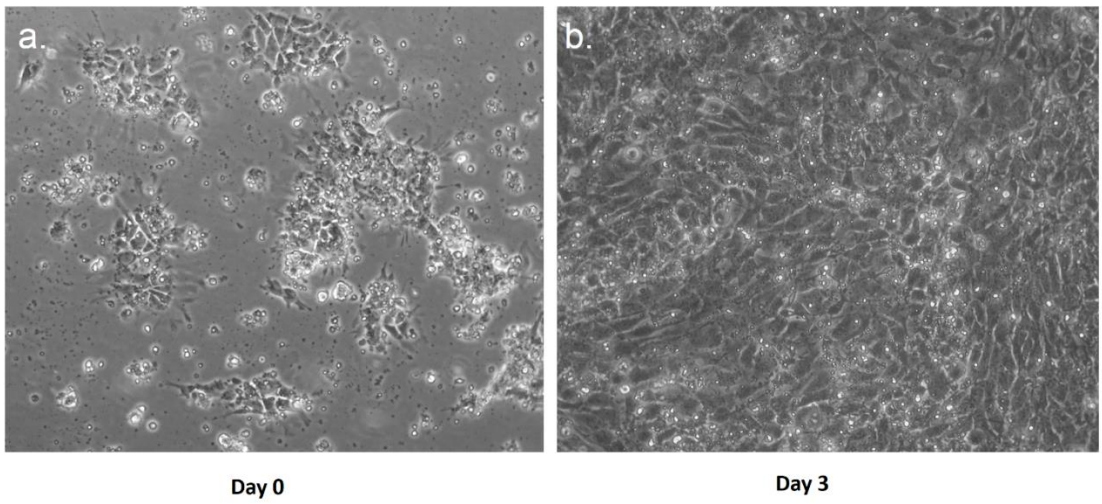


Figure 4-2. Morphology of expanded endodermal cells. a) Cells after a few hours of plating using phase microscopy at 20x. b) Endodermal cells after 3 days of expansion at 20x. Scale bar. 50um

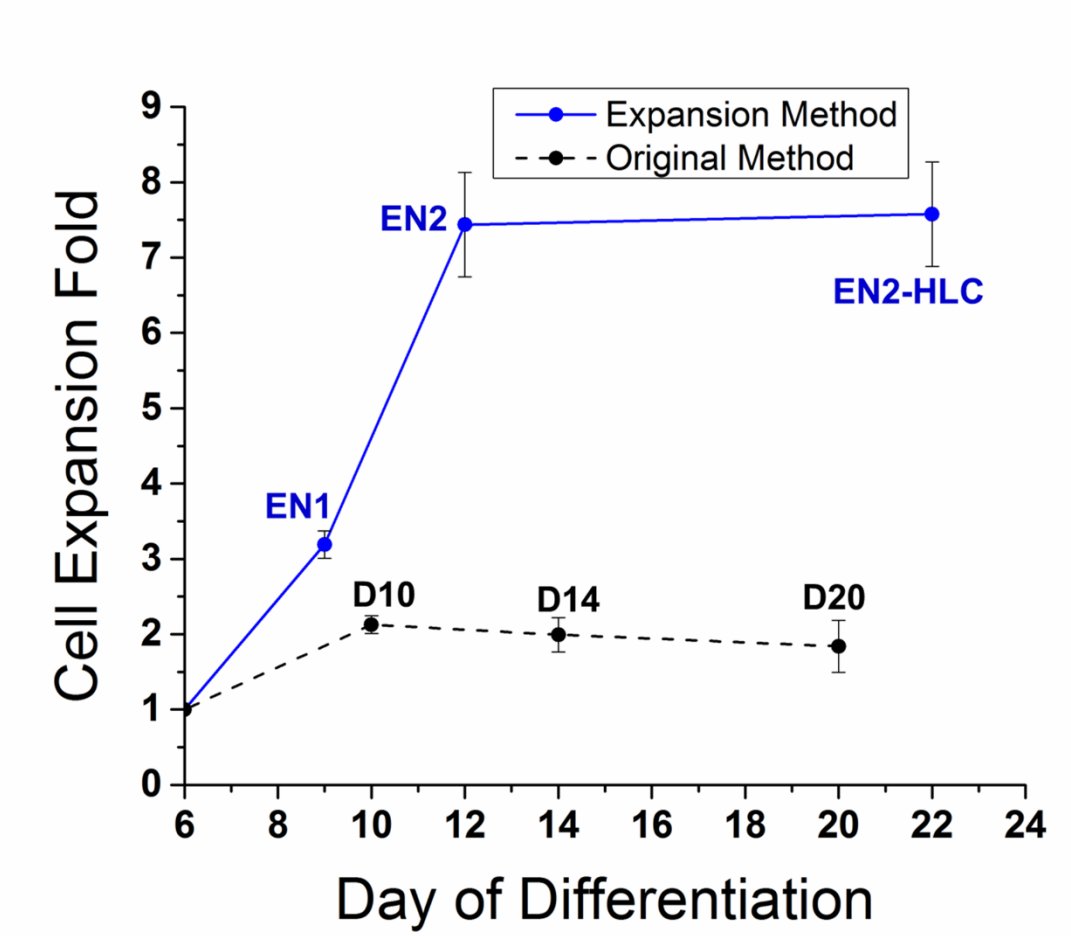


Figure 4-3. Cell Count during Differentiation. Increasing cell number during the course of cell expansion and regular differentiation. All error bars are represented as SD (n=4). SD= standard deviation

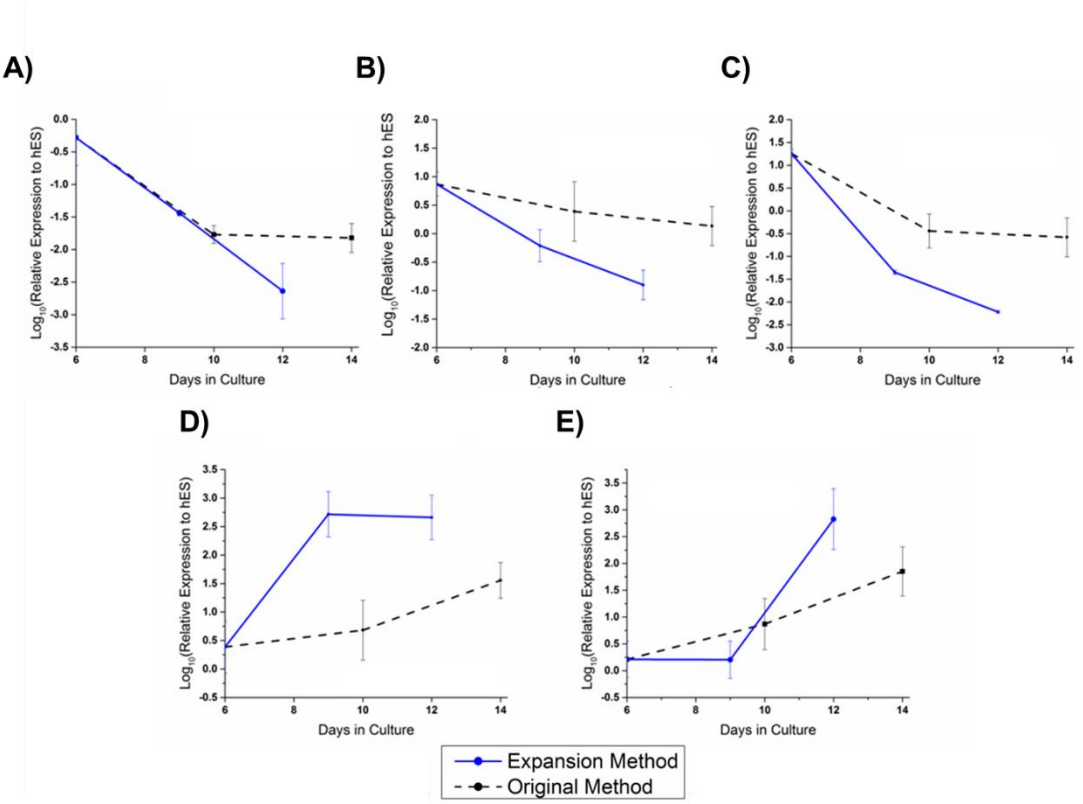


Figure 4-4. Phenotype of Endodermal Cells During Expansion. Transcript level of marker genes A) Oct4, B) CXCR4, C)GSC, D) AFP, E) ALB in endodermal cells and their differentiated progenies for the dashed black line) conventional method of differentiation or in the expansion method.

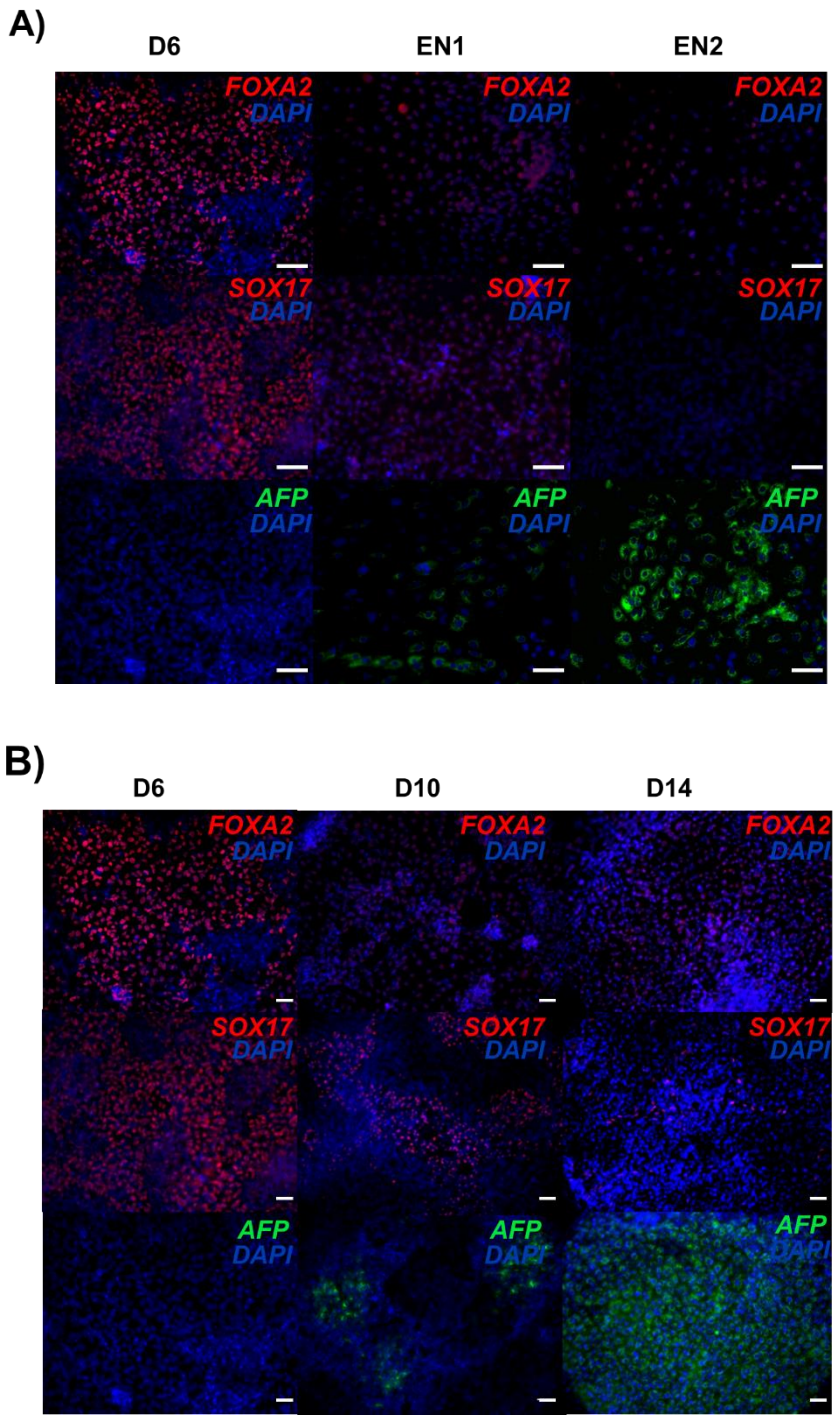


Figure 4-5. Protein expression in endodermal cells over the course of differentiation and expansion. A) Immunostaining of endodermal markers (*FOXA2 and SOX17*) and hepatic marker (AFP) during the expansion process showing simultaneous differentiation and expansion. B) Immunostaining of endodermal markers (*FOXA2 and SOX17*) and

hepatic marker (*AFP*) during the differentiation process showing differentiation occurring.

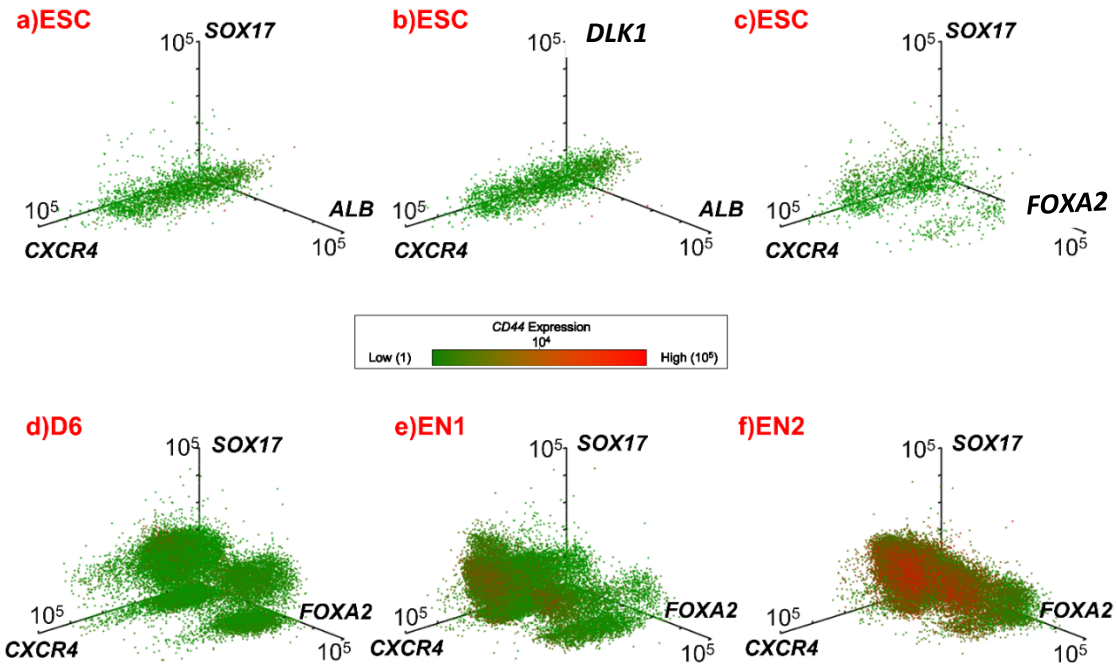


Figure 4-6. 3D plots of Early Markers During Expansion Using Mass Cytometry. Mass cytometry data is represented as a three-dimensional scatter plot with color representing the fourth dimension. The axes show the expression of different markers FOXA2, SOX17, CXCR4, *DLK1, and ALB* and the color gradient of each dot show the expression of CD44. The negative control for the markers used was hESCs which shows low expression for hepatic markers (DLK1 and ALB). As the cells undergo endodermal expansion, the endodermal genes decrease and CD44 expression increases. The presence of CD44 has been shown to play an important role in HGF-induced signaling by promoting autophosphorylation of the HGF receptor, c-MET (96). Thus, CD44 is typically used as a marker for a subpopulation of hepatic progenitor cells that have high growth potential in culture (97). The protein levels of SOX17 and CXCR4, very low in ESC (a-c), began to be co-expressed in a distinct population (Figure 4-6d). During the endodermal expansion, a

very large fraction of cells co-expressing SOX17 and CXCR4 gradually showed increased expression of CD44 (d-f).

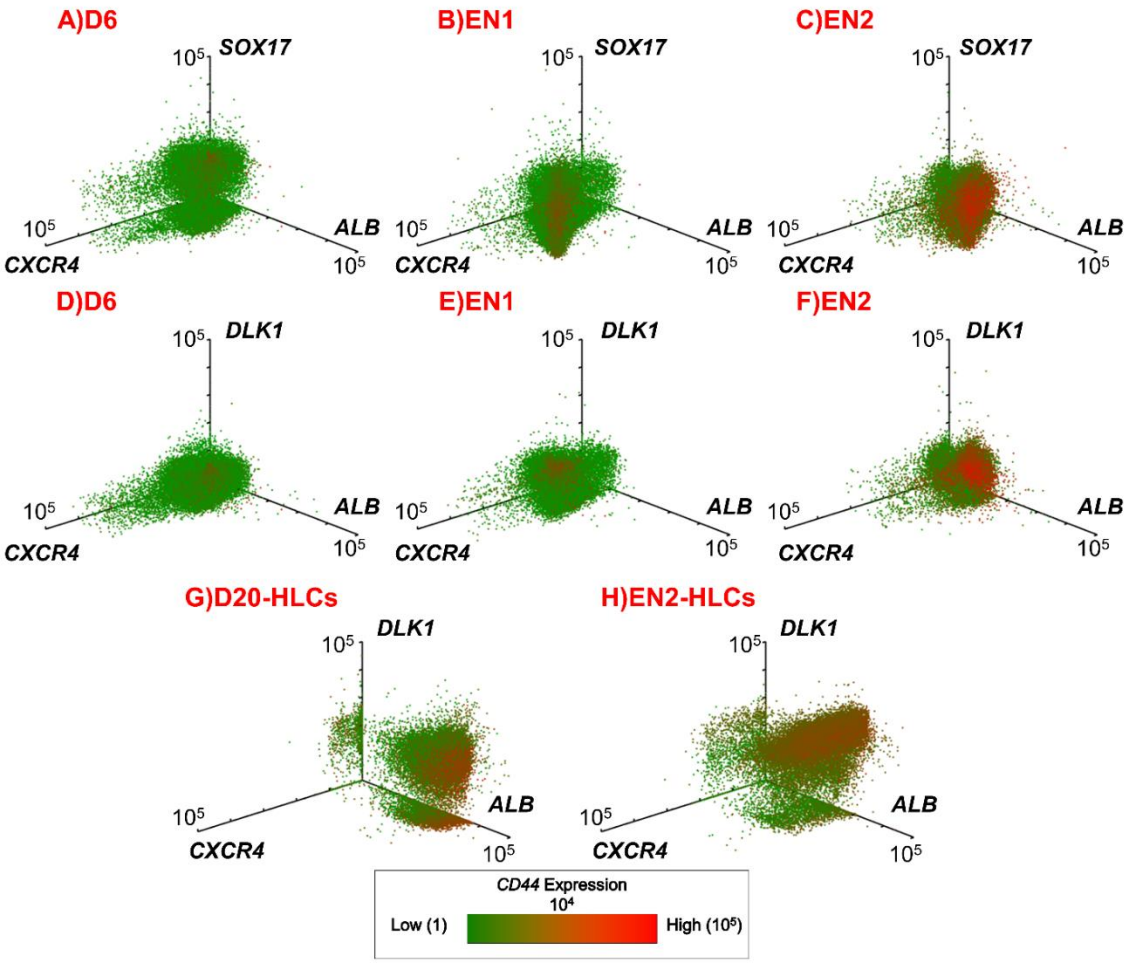


Figure 4-7. 3D plots of Late Markers During Differentiation Using Mass Cytometry. Mass cytometry data is represented as a three-dimensional scatter plot with color representing the fourth dimension. The axes show the expression of different hepatic and endoderm markers ALB, SOX17, CXCR4, or DLK1 and the colored dots show the expression of CD44. As the cells undergo endodermal expansion, the endodermal genes decrease and CD44 expression increases. Evaluation of the expanding D6 population illustrated that the bulk of the cell population expressing both ALB and CD44 also

simultaneously lose expression of the endodermal markers, SOX17 and CXCR4, in line with results from qRT-PCR and immunostaining. DLK1 has frequently been used as a marker for early hepatocyte committed progenitor cells that can be isolated from the embryonic liver *in vivo*(112, 113) (74, 75). Using DLK1 as a marker for hepatocyte-committed cells, a large fraction of the expanding EN1 and EN2 population began to gain DLK1 expression giving credence to the assignment of early hepatic subpopulation. The expression of CD44 and ALB were gradually increasing along with DLK1, showing co-expression of all three hepatoblast markers in the majority of DLK1 positive cells. The co-expression of hepatoblast markers observed in the D20-HLCs and EN2-HLCs populations further illustrates that the majority of the cells are capable of differentiating towards the hepatocyte lineage regardless of whether or not they were subjected to endodermal expansion steps. Notably both D20-HLCs and EN2-HLCs showed low expression levels of CXCR4.

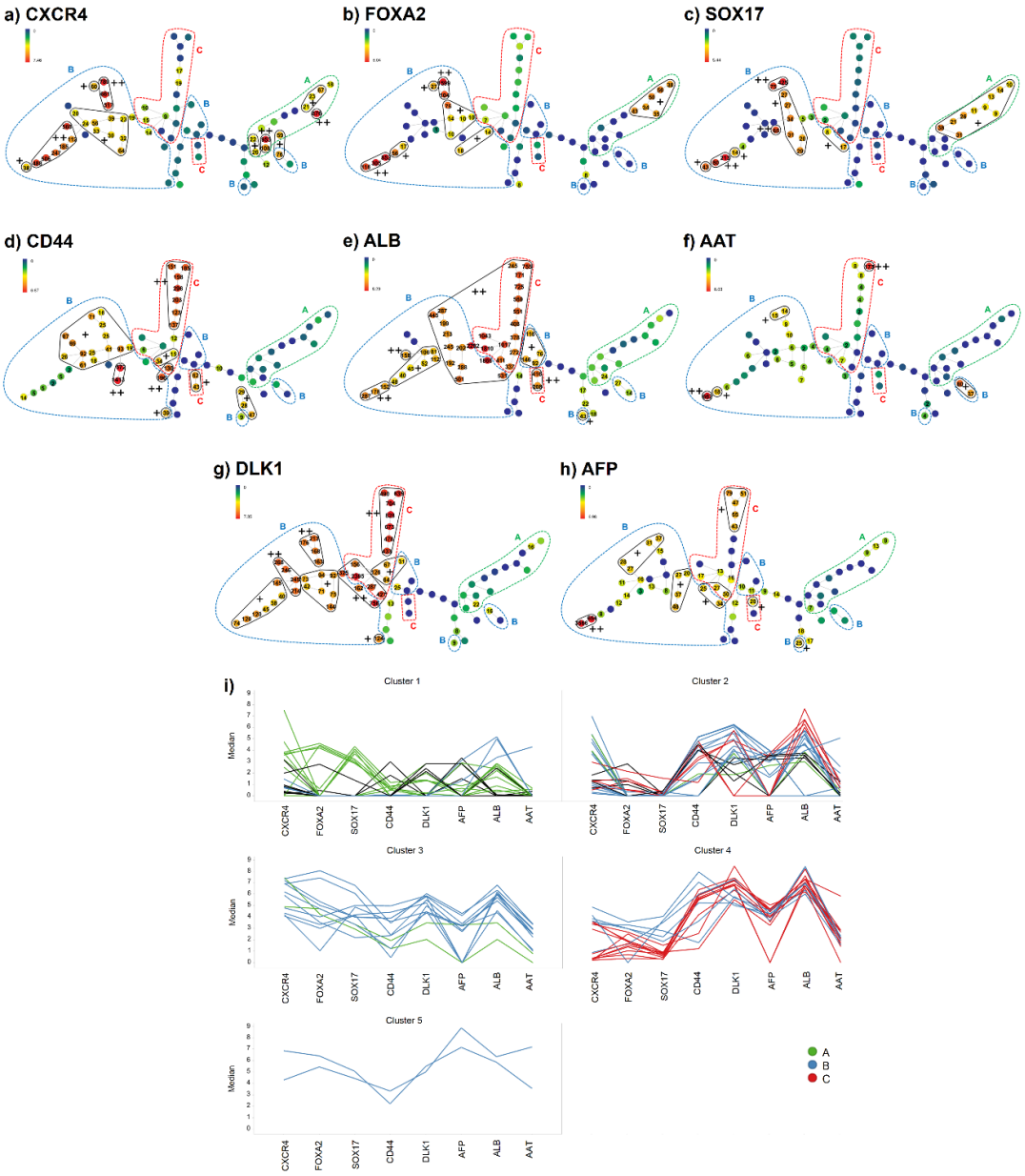


Figure 4-8. SPADE analysis of Mass Cytometry Data. (a-h) The expression level of each of 8 markers is shown in a separate tree diagram. The color represents arcsinh transformed intensity of each marker. Median of the intensity for a node is shown. The expression level of each marker is ranked as high (++), moderate (+), and low (-). The threshold for ++ and + is 100 and 20 for CXCR4, 100 and 10 for FOXA2, 60 and 8 for

SOX17, 100 and 15 for CD44, 150 and 40 for ALB, 100 and 14 for AAT, 150 and 30 for DLK1, and 100 and 20 for AFP, respectively. Unannotated parts of the SPADE tree for each marker were considered as the cells with low (-) expression of the marker. i) K-means clustering was performed to the 75 nodes from the SPADE analysis based on the median intensity of the 8 markers (green lines: nodes in subpopulation A (endodermal cells), blue lines: nodes in subpopulation B (hepatic endoderm/hepatoblasts), red lines: nodes in subpopulation C (hepatoblasts/hepatocytes), and black lines: nodes not classified as any subpopulation). The distance was measured with Euclidean distance. 10 of the 14 nodes in the A subpopulation were in Cluster 1. 11 of the 16 nodes in the C subpopulation were in Cluster 4, revealing that majority of the nodes in A and C were separately clustered. The majority of the B nodes were interspersed into Clusters 2 and 3, with 12 and 10 of the B nodes in Clusters 2 and 3, respectively.

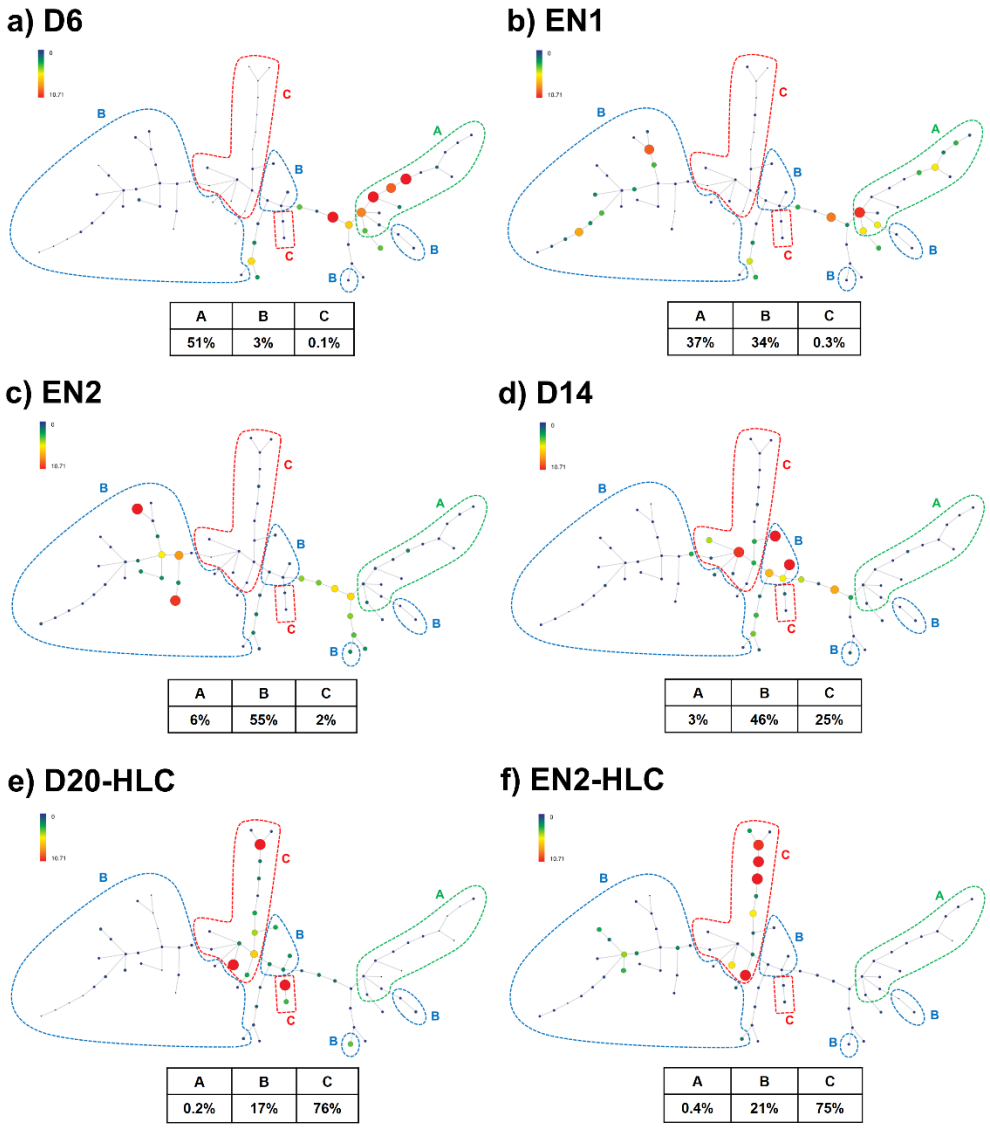


Figure 4-9. SPADE Analysis of Mass Cytometry Data of Cells Differentiating toward HLCs. The mass cytometry data of cells at different differentiation stages were clustered into 75 nodes by SPADE analysis. SPADE results for (a) D6, (b) EN1, (c) EN2, (d) D14, (e) D20-HLC, and (f) EN2-HLC are shown. The color and the size of a node represent percentage of the populations for the node in each sample. Subpopulations enclosed in A, B, and C are classified as endodermal cells, hepatic endoderm/hepatoblasts, and hepatoblasts/hepatocytes, respectively. The percentage of endodermal cells, hepatic endoderm/hepatoblasts, and hepatoblasts/hepatocytes for each population is shown.

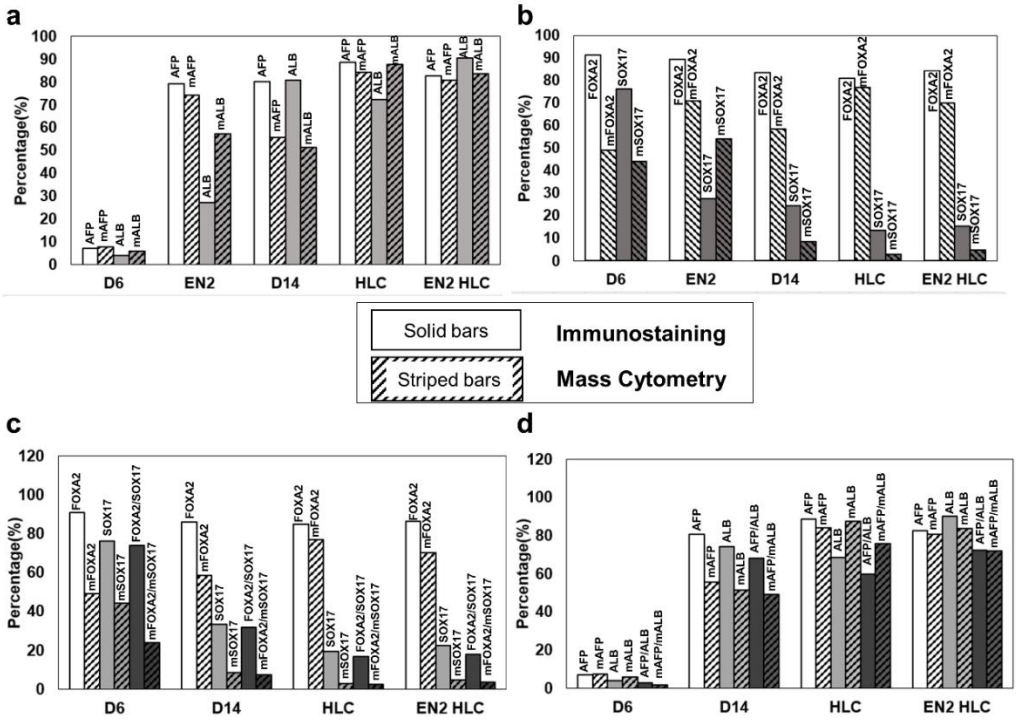


Figure 4-10. Mass cytometry and immunostaining of hepatic and endodermal markers. a) Immunostaining for AFP (white) and ALB (grey) was carried out at different stages of the differentiation along with the percentage of cells positive for AFP (white with lines) and ALB (grey with lines) found during mass cytometry. b) Immunostaining for FOXA2 (white) and SOX17 (grey) was carried out at different stages of the differentiation along with the percentage of cells positive for FOXA2 (white with lines) and SOX17 (grey with lines) found during mass cytometry. c) Co-staining was carried out for FOXA2 (white) and SOX17 (grey) where the percentages of double positive (dark grey) cells were quantified and also compared to mass cytometry percentages (stripe bars). d) Co-staining was carried out for AFP (white) and ALB (grey) where the percentages of double positive (dark grey) cells were quantified and also compared to mass cytometry percentages (stripe bars).

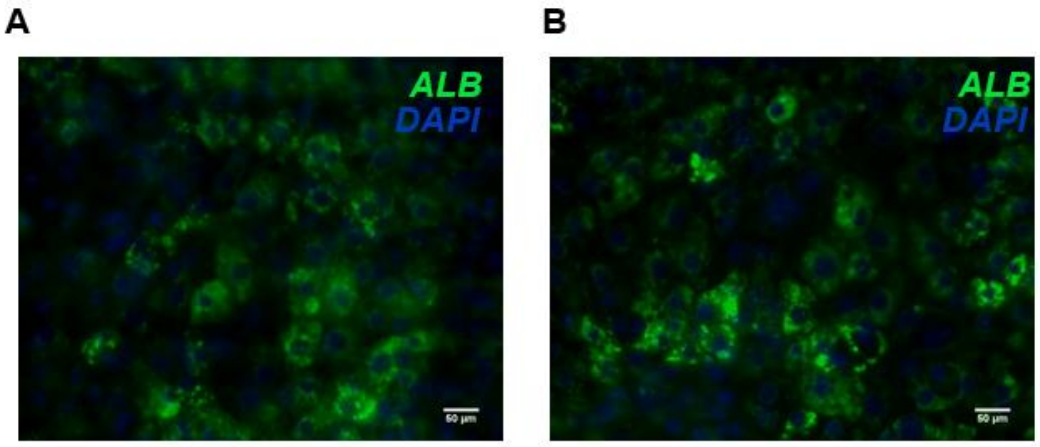


Figure 4-11. Immunostaining of Albumin in HLCs. in A) HLCs and B) EN2-HLCs. Scale bar=50um.

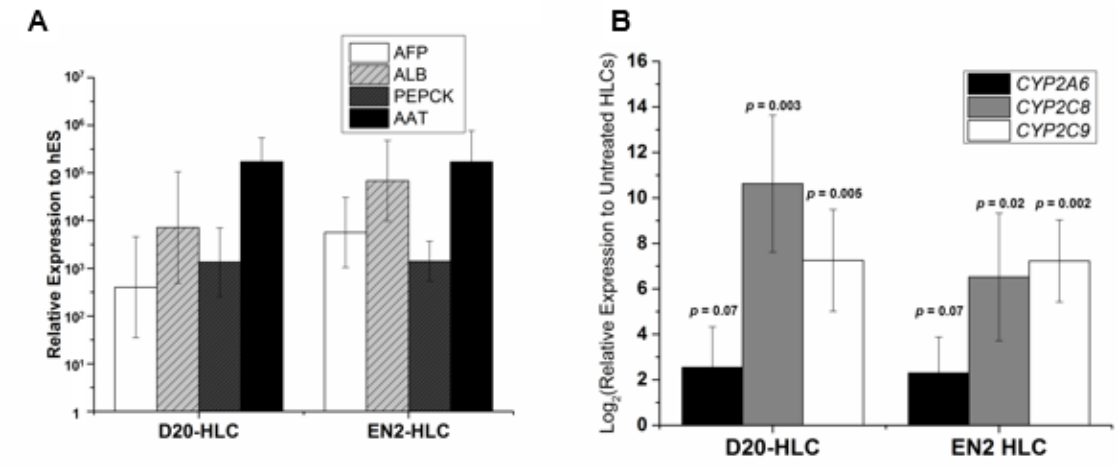


Figure 4-12. Transcript Level of Different Hepatic Markers in HLCs. A) D20-HLCs and EN2-HLCs both expressed transcripts of hepatocyte markers AFP, ALB, AAT, and PEPCK. B) Transcript level change of CYP2A6, CYP2C8, and CYP2C9 in D20-HLCs and EN2-HLCs after induction with rifampicin. P values of Student's t-test are shown. n=3

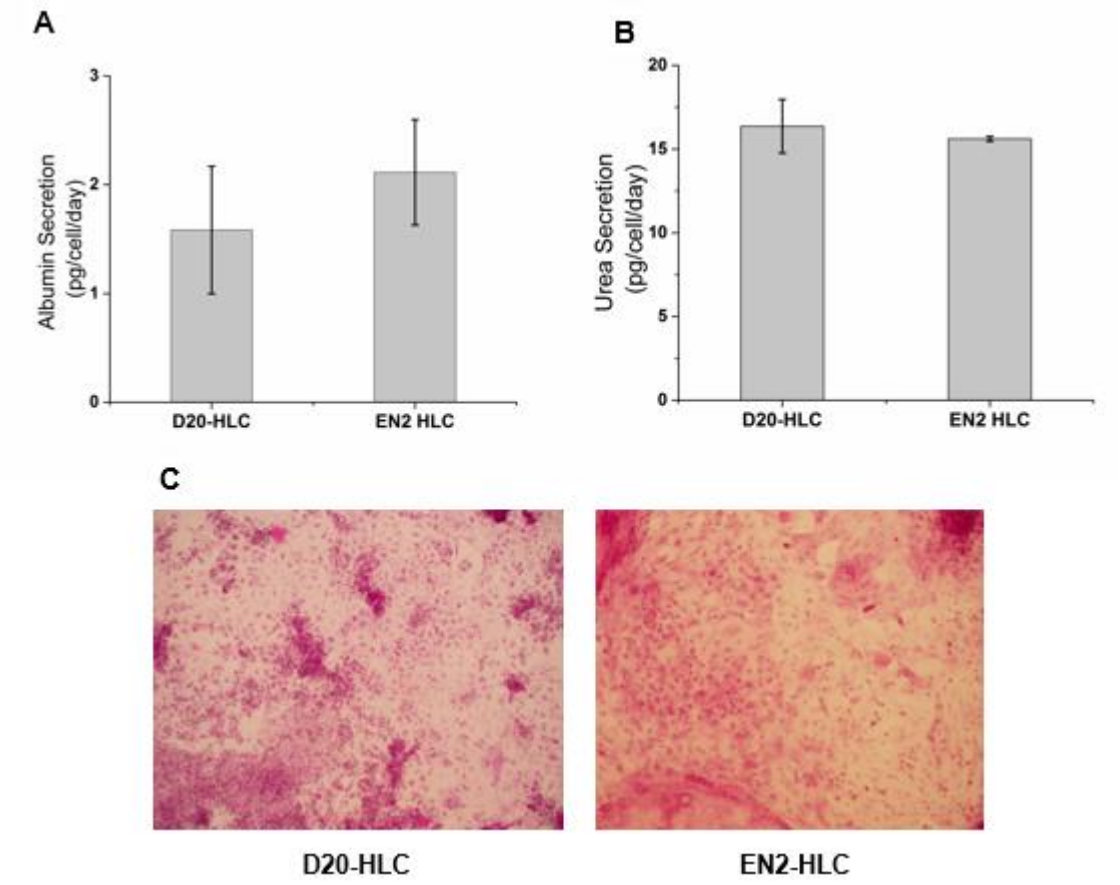


Figure 4-13. Functional Activity of HLCs and EN2-HLCs. A) Albumin secretion for D20-HLCs and EN2-HLCs B) Urea Secretion for D20-HLCs and EN2-HLCs C) Glycogen staining for D20-HLCs and EN2-HLCs. All error bars are represented as SD. n=3

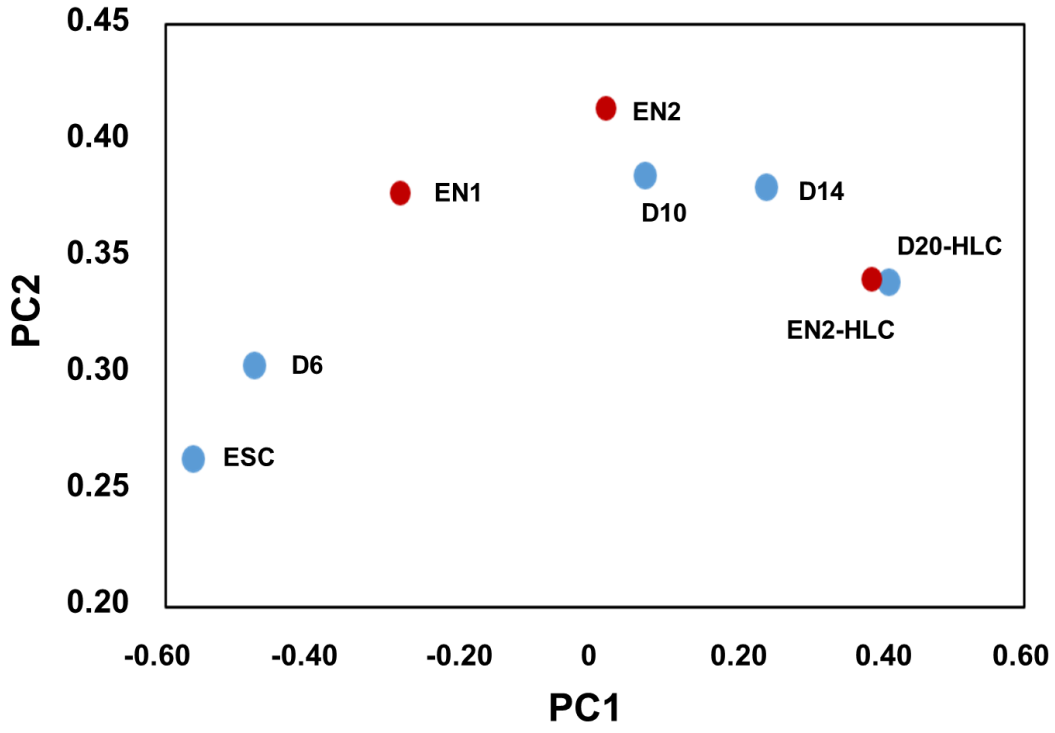


Figure 4-14. Transcriptome analysis of differentiating and expanding cells. Cell samples at different differentiation stages from the conventional protocol (blue) and expanding protocol (red) align along the same PC space (PC1 vs. PC2). EN1 and EN2 samples are in between D6 and D10 (the beginning and end of Stage 2 differentiation of the conventional protocol). PC, principal component.

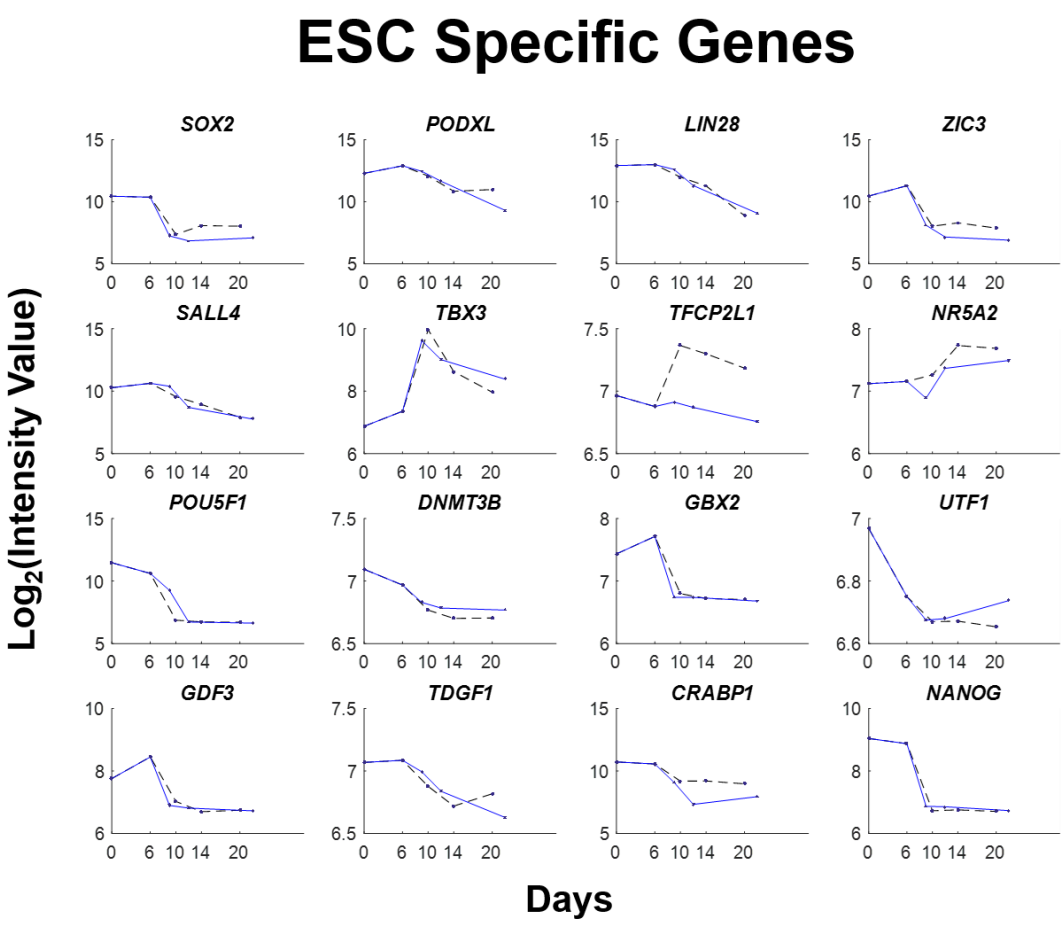


Figure 4-15. Transcript dynamics of embryonic stem cell specific genes during directed differentiation with the original protocol or with simultaneous expansion.

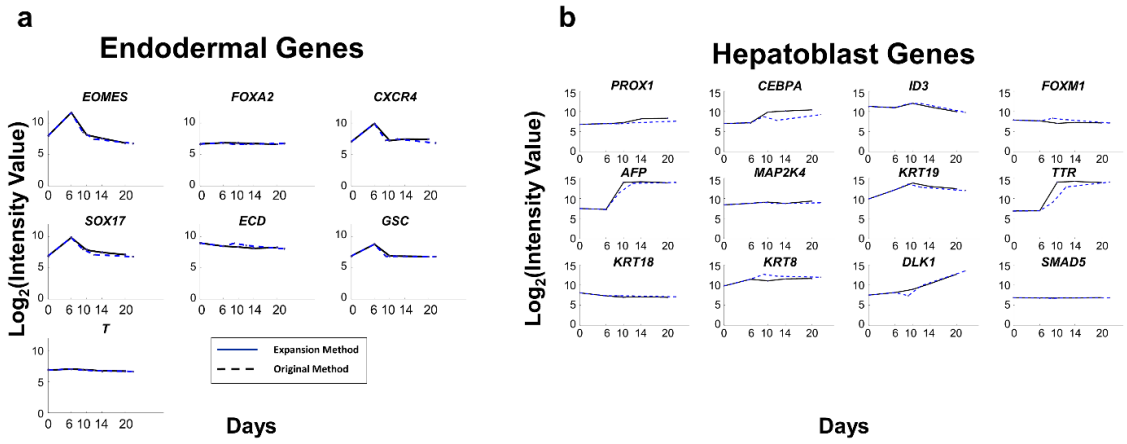
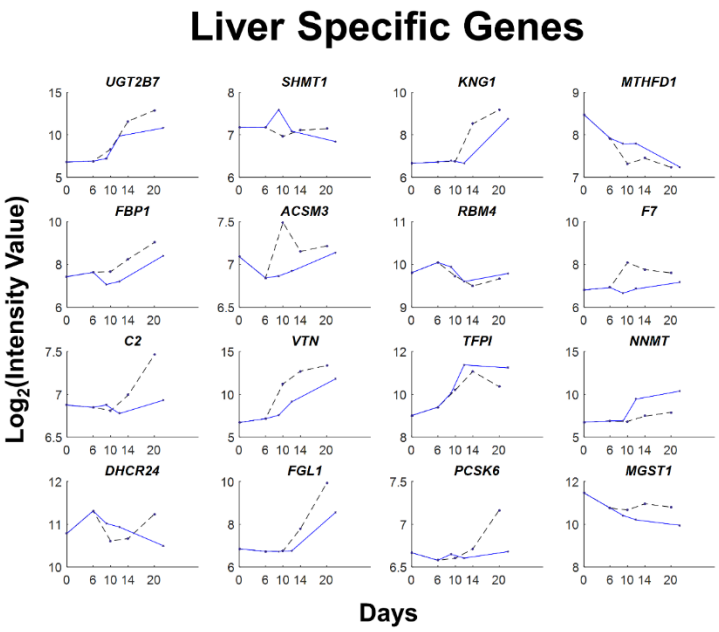
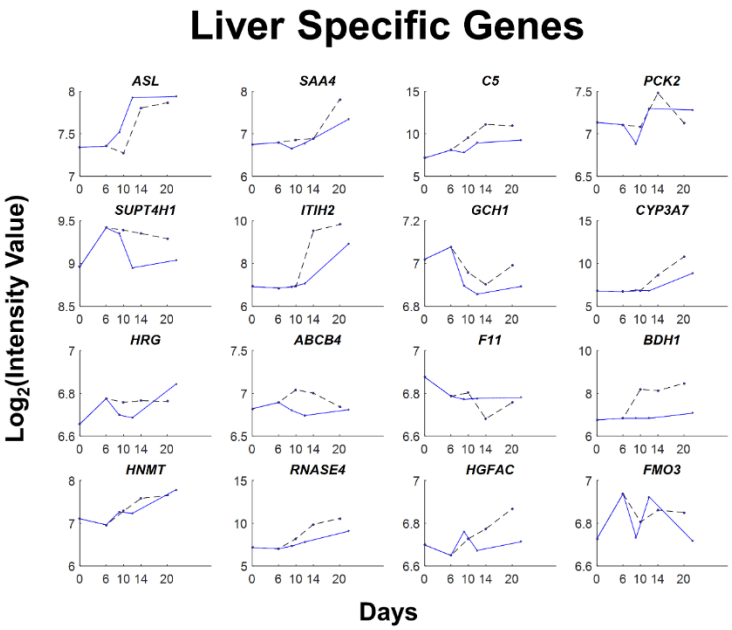


Figure 4-16. Transcript levels of stage-specific markers. Dynamics of key a) endoderm markers and b) hepatoblast markers is similar in cells differentiated with the two different protocols (dashed = original method, solid = expansion method).

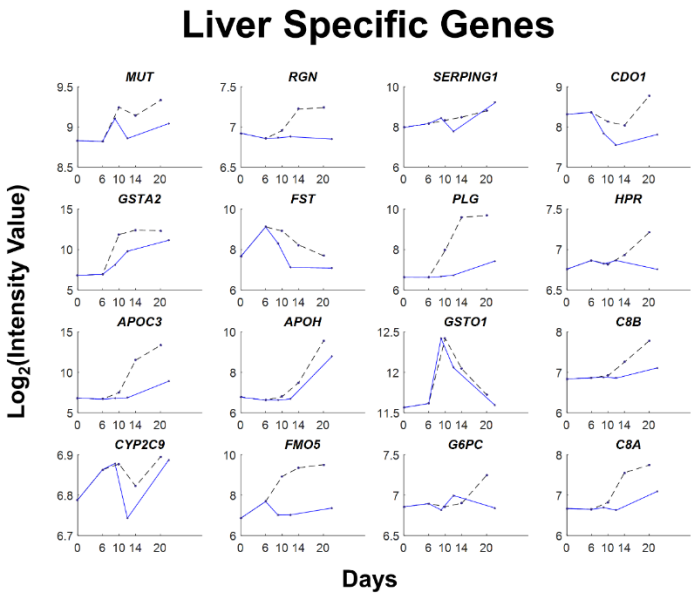
a


Liver Specific Genes
UGT2B7
SHMT1
KNG1
MTHFD1
FBP1
ACSM3
RBM4
F7
C2
VTN
TFPI
NNMT
DHCR24
FGL1
PCSK6
MGST1
Log2(Intensity Value)
Days

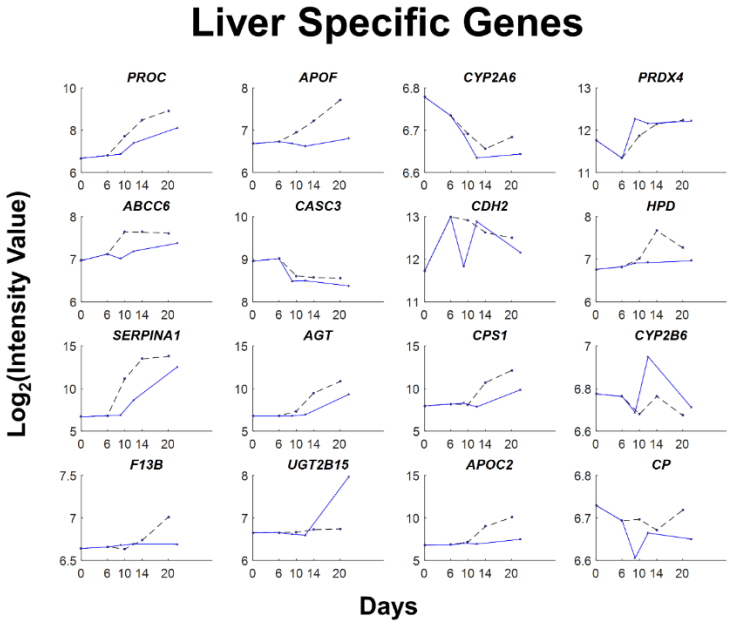
b


Liver Specific Genes
ASL
SAA4
C5
PCK2
SUPT4H1
ITIH2
GCH1
CYP3A7
HRG
ABCB4
F11
BDH1
HNMT
RNASE4
HGFAC
FMO3
Log2(Intensity Value)
Days

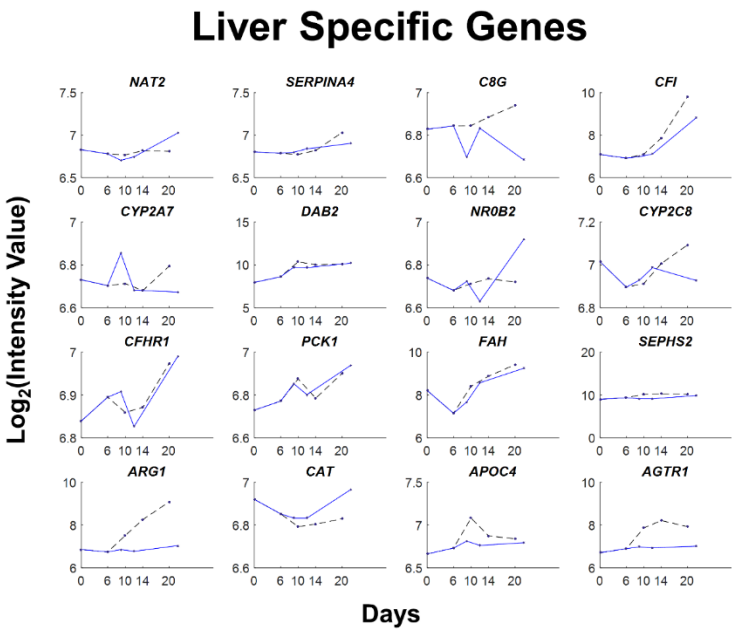
c


Liver Specific Genes
MUT
RGN
SERPING1
CDO1
GSTA2
FST
PLG
HPR
APOC3
APOH
GSTO1
C8B
CYP2C9
FMO5
G6PC
C8A
Log2(Intensity Value)
Days

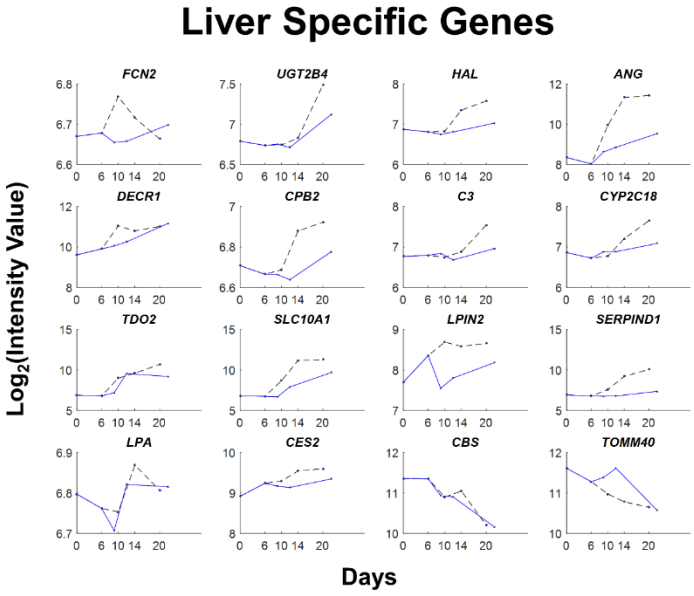
d


Liver Specific Genes
PROC
APOF
CYP2A6
PRDX4
ABCC6
CASC3
CDH2
HPD
SERPINA1
AGT
CPS1
CYP2B6
F13B
UGT2B15
APOC2
CP
Log2(Intensity Value)
Days

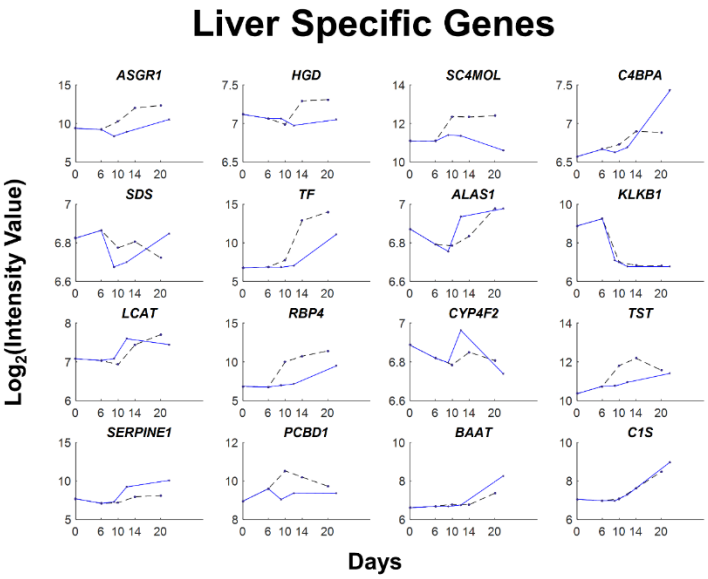
e


Liver Specific Genes
NAT2
SERPINA4
C8G
CFI
CYP2A7
DAB2
NR0B2
CYP2C8
CFHR1
PCK1
FAH
SEPHS2
ARG1
CAT
APOC4
AGTR1
Log2(Intensity Value)
Days

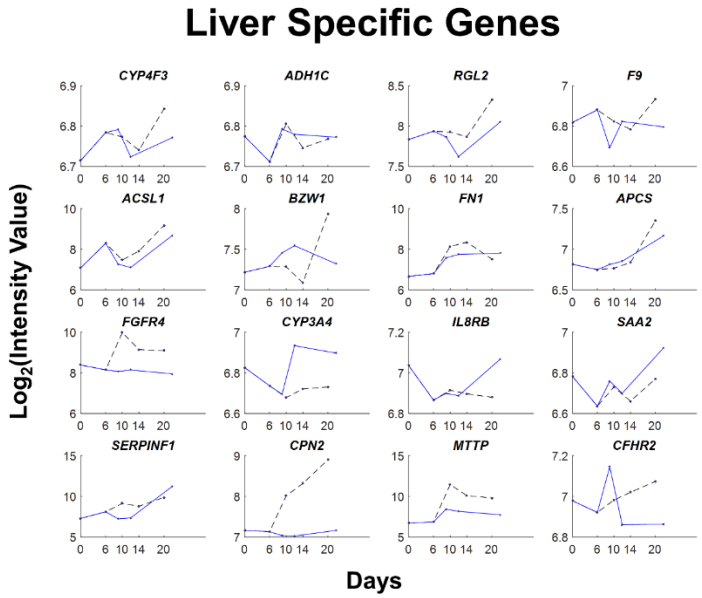
f


Liver Specific Genes
FCN2
UGT2B4
HAL
ANG
DECR1
CPB2
C3
CYP2C18
TDO2
SLC10A1
LPIN2
SERPIND1
LPA
CES2
CBS
TOMM40
Log2(Intensity Value)
Days

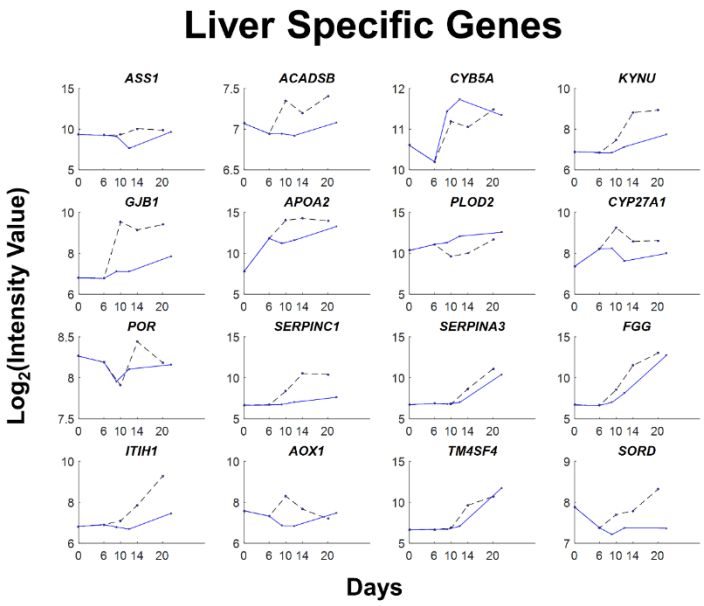
g


Liver Specific Genes
ASGR1
HGD
SC4MOL
C4BPA
SDS
TF
ALAS1
KLKB1
LCAT
RBP4
CYP4F2
TST
SERPINE1
PCBD1
BAAT
C1S
Log2(Intensity Value)
Days

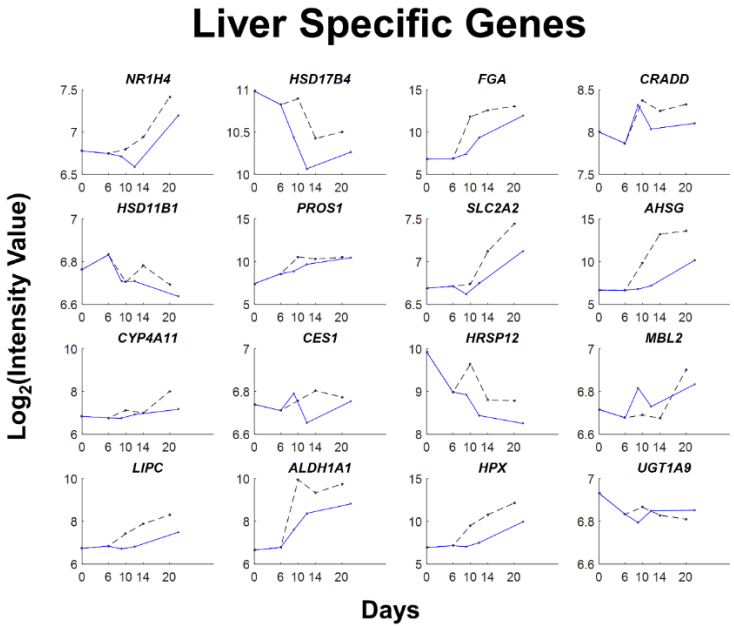
h


Liver Specific Genes
CYP4F3
ADH1C
RGL2
F9
ACSL1
BZW1
FN1
APCS
FGFR4
CYP3A4
IL8RB
SAA2
SERPINF1
CPN2
MTTP
CFHR2
Log2(Intensity Value)
Days

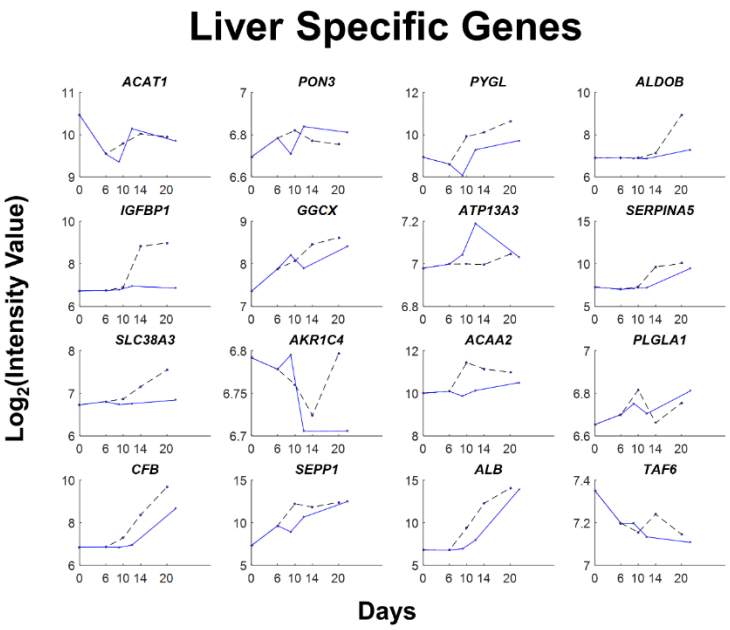
i


Liver Specific Genes
ASS1
ACADSB
CYB5A
KYNU
GJB1
APOA2
PLOD2
CYP27A1
POR
SERPINC1
SERPINA3
FGG
ITIH1
AOX1
TM4SF4
SORD
Log2(Intensity Value)
Days

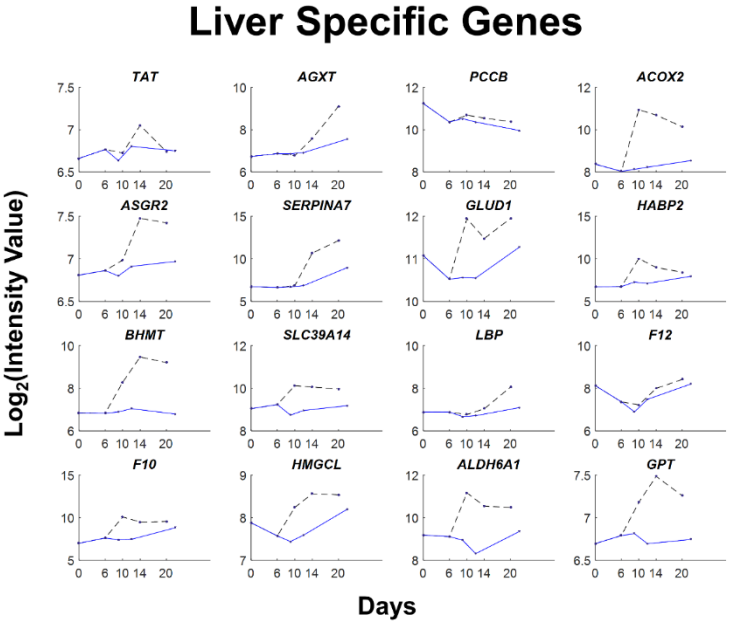
j


Liver Specific Genes
NR1H4
HSD17B4
FGA
CRADD
HSD11B1
PROS1
SLC2A2
AHSG
CYP4A11
CES1
HRSP12
MBL2
LIPC
ALDH1A1
HPX
UGT1A9
Log2(Intensity Value)
Days

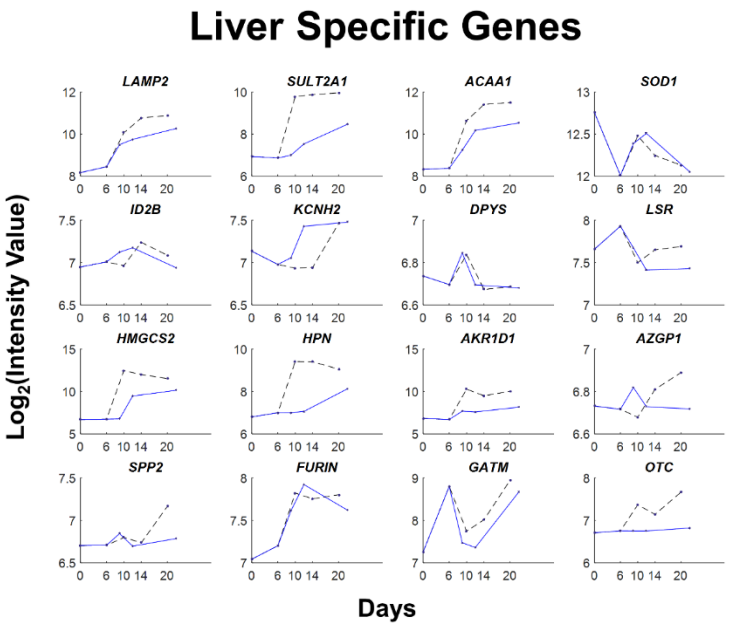
k


Liver Specific Genes
ACAT1
PON3
PYGL
ALDOB
IGFBP1
GGCX
ATP13A3
SERPINA5
SLC38A3
AKR1C4
ACAA2
PLGLA1
CFB
SEPP1
ALB
TAF6
Log2(Intensity Value)
Days

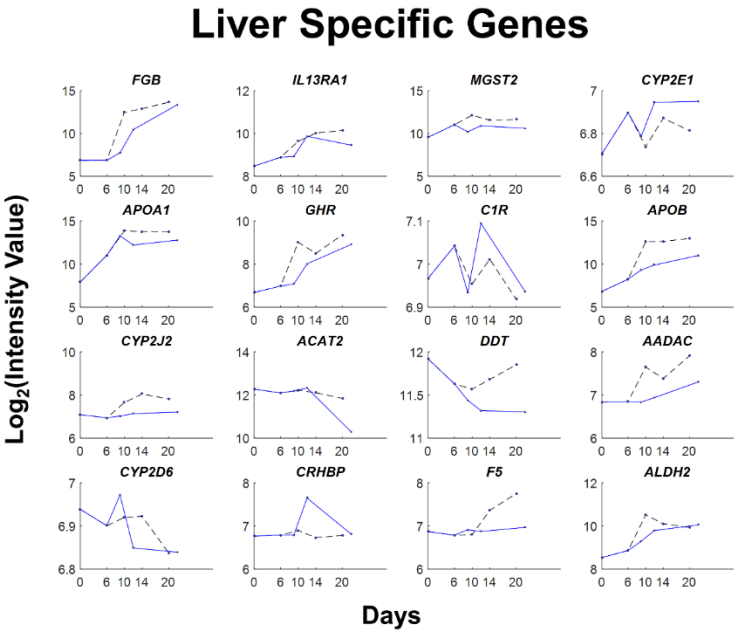
l


Liver Specific Genes
TAT
AGXT
PCCB
ACOX2
ASGR2
SERPINA7
GLUD1
HABP2
BHMT
SLC39A14
LBP
F12
F10
HMGCL
ALDH6A1
GPT
Log2(Intensity Value)
Days

m


Liver Specific Genes
LAMP2
SULT2A1
ACAA1
SOD1
ID2B
KCNH2
DPYS
LSR
HMGCS2
HPN
AKR1D1
AZGP1
SPP2
FURIN
GATM
OTC
Log2(Intensity Value)
Days

n


Liver Specific Genes
FGB
IL13RA1
MGST2
CYP2E1
APOA1
GHR
C1R
APOB
CYP2J2
ACAT2
DDT
AADAC
CYP2D6
CRHBP
F5
ALDH2
Log2(Intensity Value)
Days

o

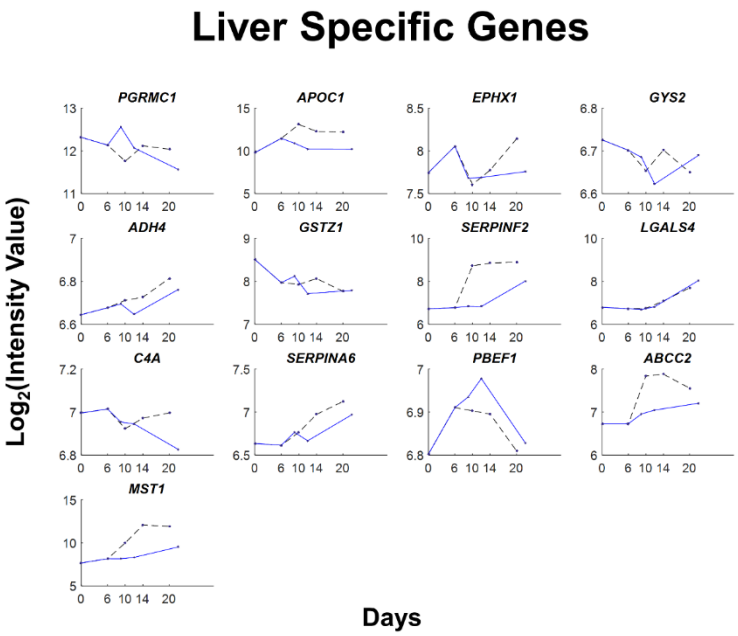


Figure 4-17a-o. Transcript dynamics of embryonic liver specific genes during directed differentiation with the original protocol or with simultaneous expansion.

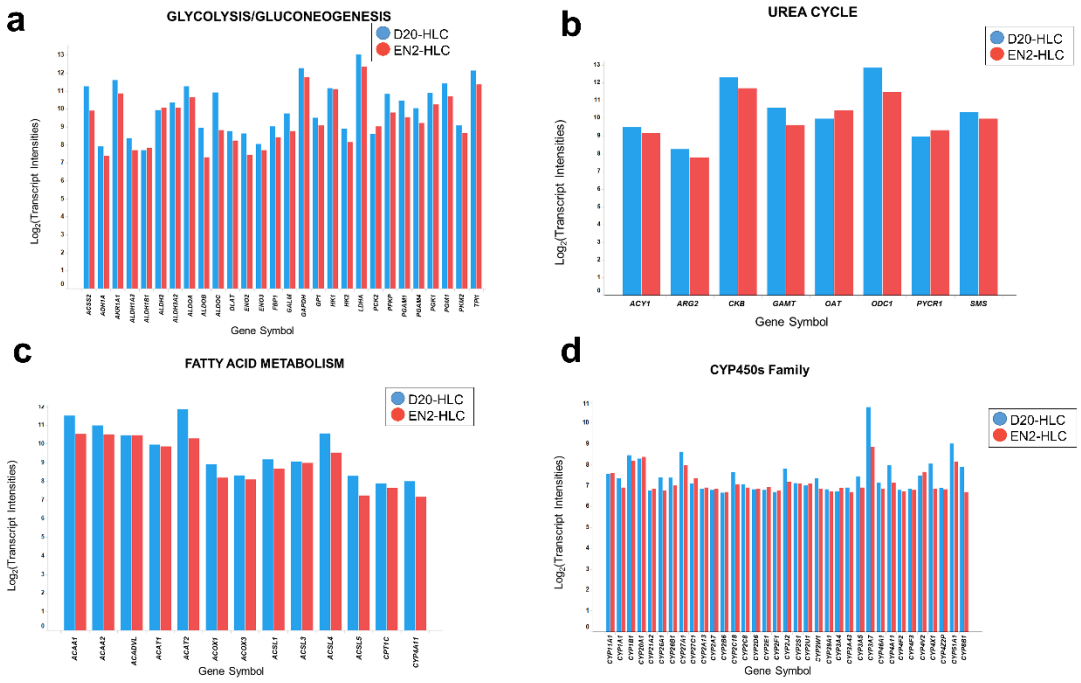


Figure 4-18. Comparison of transcript levels of genes of different functional classes between D20-HLC and EN2-HLC. a) Expression levels of glycolysis/gluconeogenesis genes were similar between D20-HLCs and EN2-HLCs. b) Expression levels of urea cycle genes were similar between D20-HLCs and EN2-HLCs. c) Expression levels of fatty acid metabolism genes were similar between D20-HLCs and EN2-HLCs. d) Expression levels of genes in the cytochrome p450 family were similar between D20-HLCs and EN2-HLCs.

Chapter 5. Transcriptome Analysis of Stem Cell Differentiation to Hepatic Lineage

5.1 Introduction

In the previous chapter, we demonstrated that hESCs can be differentiated towards the hepatic lineage and result in cells referred to as HLCs. These cells show remarkable similarities to primary hepatocytes that have the potential to be used in different clinical and industrial applications. They have increased hepatic-specific gene expression, protein expression and function similar to primary hepatocytes. However, when we compare our functional characterization to studies on primary hepatocytes, it is clear that directed differentiation from hESCs only give us immature phenotype, hence the name "hepatocyte-like". Although the field of developmental biology has come a long way and have continued to provide us with insights into different signaling pathways that have allowed us to develop our differentiation protocol, there is still a roadblock that is preventing us from achieving a primary hepatocyte state.

Many approaches have been explored to enhance maturation of HLCs, including the formation of tissue-like 3D structure (102) (114), co-culture of HLCs and endothelial cells (114-116), the addition of small molecules uncovered by screening (117-119) , and overexpression of certain transcription factors(120-122). Using cell aggregates formed by co-culturing HLCs and endothelial cells for transplantation, some mature liver markers were seen in the transplanted cells (114). However, despite the improvements, HLCs still remain functionally immature when compared to their primary counterparts in liver. The general lack of maturity of HLCs was also revealed in a microarray transcriptome study showing that HLCs derived from three different laboratories were clustered distinctively from primary hepatocytes (123). Through functional activities and proteomic analysis, it was clearly demonstrated that HLCs were closer to fetal liver hepatocytes than primary hepatocytes (123, 124).

In this chapter, we will be trying to understand the universal roadblock that is preventing the immature phenotype by looking at whole gene expression analysis. This will be accomplished by using transcriptome analysis which is very similar to qPCR in concept and sample preparation except we can look at it from a whole transcriptome point of view. Using this approach we hope to see a more holistic view of our differentiation to understand what signaling pathways are missing or need to be activated. Total RNA lysate is obtained from the samples followed by purification of the mRNA fraction similar to prepping for qPCR. From this mRNA lystate, a complementary DNA is generated which can then be amplified using PCR and hybridized to the microarray chip. Although there are many technologies that are available for transcriptome analysis, we are focused on using microarray because of the cost and ease of data processing. In this system, we can look at over 47,000 probes that cover over 27,000 genes.

Using whole transcriptome analysis, we can identify certain genes or signaling pathways that are not expressed correctly compared to primary hepatocytes. Although the lack of maturity of HLCs has been well accepted, it is not known at which stage of embryonic liver development HLCs resemble. Comparison of transcriptome data of HLCs with those from embryonic development should allow us to identify the equivalent embryonic liver developmental stage for HLC. This may also lead to the identification of genes that behave "erratically" in HLCs differentiation compared to embryonic liver development. This discovery may potentially reveal the genetic roadblocks preventing HLCs from reaching the functional maturity of their tissue counterparts.

Human gestation occurs over 280 days compared to 20 days in mouse. A typical directed stem cell differentiation process for both human and mouse PSCs lasts for about 20 days (88). Although these processes differ in time scales, they share similar progression through the different developmental stages. A key to the cross-species and cross- *in vitro*-*in vivo* analysis is to align all the different data onto a common "time scale". We hypothesize that the transcriptome data of *in vitro* hepatic differentiation bear the characteristics of different liver embryonic stages in human and mouse. While transcriptome data of *in vitro* hPSC differentiation to HLC are publicly available, human *in vivo* data at various stages of liver development are difficult to obtain. On the other hand,

transcriptome data of *in vivo* rodent liver development are more publicly accessible. We hypothesized that the gene expression dynamics of embryonic liver development between rodents and human are highly similar, and set out to identify the equivalent stage of rodent embryonic liver development for human HLCs by a cross-species transcriptome meta-analysis.

Systematic variations and random errors in the transcriptome data from different studies can give biasing that resemble the characteristics of the sample source. Samples assayed with different measurement platforms can also bear the characteristics of the platform. The sample source and assay platform derived-variations that can arise need to be removed before the transcriptome data can be combined and analyzed. The ComBat algorithm is an empirical Bayes method that performs a location and scale adjustment by pooling information across each gene in every batch and using this information to center the data to an overall grand mean (125). This technique can be applied to multiple batches of datasets and is robust for small sample size. Other methods of batch correction also have been described in detail such as the *limma* package which tends to give less "exaggerated" effects (126). ComBat consistently outperformed the common batch correction techniques, including distance-weighted discrimination (DWD) which uses a support vector machine based approach (127)(127), mean-centering (PAMR) which relies on gene based analysis of variance (128), and surrogate variable analysis (SVA) which uses a combination of single value decomposition and linear model analysis(129, 130).

In this study, we employed the ComBat algorithm to integrate transcriptome data from *in vitro* differentiation of h/mPSCs and *in vivo* mouse embryonic liver development obtained from different studies and assay platforms. The batch correction allows the data of mouse embryonic liver development and HLCs derived from h/mPSCs to be aligned according to a common scale of development. Past studies have relied on the comparison of gene expression profile of primary hepatocytes to HLCs obtained at the end of differentiation to identify differentially expressed genes(124, 131)(86, 93). By placing the gene expression profile from the meta-data on a common developmental scale, we can use expression dynamics to more accurately identify genes that deviate from embryonic liver

development compared to *in-vitro* differentiation. These genes might contribute to the lack of maturity in HLCs differentiated from PSCs. The described meta-analysis of the highly heterogeneous transcriptome data from human and mouse, and *in vivo* and *in vitro* data, can provide clues for genetic intervention to advance directed differentiation of PSCs to hepatocytes, and be used for possibly other lineage differentiations.

5.2 Approach to Transcriptome Meta-Analysis

Transcriptome data was collected from multiple studies to increase the validity and complexity of our dataset. We also included transcriptome data of our differentiation data from multiple cell lines from human and mouse cells to the hepatic lineage. Our focus was on studying the roadblocks behind our differentiation process and we incorporated other laboratory transcriptome data to support our approach and prove that it is robust and can be applied to other studies as well.

5.2.1 Transcriptome Dataset

The total RNA was extracted at different time points from human and mouse liver differentiation samples using the RNeasy Mini kit (Qiagen). Differentiating human embryonic stem cells (hESCs) (H9, HSF6) samples were hybridized to the Illumina HT12 bead array v3 (Illumina Inc.), HES3 samples were hybridized to Human Genome U133 v2 array (Affymetrix) and the mouse iPSC-derived samples were hybridized to the WG-6 v2 (Illumina) array. Other datasets from different studies using different protocols were also included in this study. The source of the different human and mouse datasets is listed in Table 5-1 and 5-2, respectively. The human dataset included transcriptome of different hESCs differentiating to hepatocytes including HLCs generated using various protocols, two time course data from different protocols (Hu/Duncan) (79, 90, 132-134), HLCs in three dimensional spheroids (102), liver organoids generated through co-culture of induced pluripotent stem cells (iPSC)-derived HLCs and mesenchymal and endothelial cells (114), and iHEPs generated through direct reprogramming of fibroblasts to HLCs by hepatic transcription factor transduction(98) .Also included were primary human hepatocytes

(pooled mRNA after 1 day of culture), fetal liver (18 weeks of gestation) and adult liver. Transcriptome data of mouse embryonic liver development (E9.5 to post-natal) include $CD45^{-}Ter119^{-}$ liver cells from C57/BL6/Tg mice embryos and whole livers from C57/B6 mice embryos (114, 135).

5.2.2 Overview of Combining Datasets

Briefly, the raw data of Affymetrix, Illumina and Agilent array platforms were processed using the *affy, lumi and limm*a packages in R respectively to obtain intensity expression values. Logarithmic transformed (log_2) intensity from different platforms were corrected for depth of digital precision. The expression level of each gene is condensed by the median expression value of multiple specific probes. Common ENSEMBL identifiers were used to condense the same genes from arrays of different platforms. The cross platform data was then combined into a master expression dataset based on matching the ENSEMBL identifiers. The initial data processing until normalization was performed using the *virtualArray* package in R. Each of the human and mouse datasets were linear normalized. The average expression values from replicate samples were calculated and used for further analysis. The mouse and human datasets consisted of 16,415 genes and 17,683 genes respectively, of which 14,333 were common based on orthologous ENSEMBL identifiers.

5.2.3 Criteria for Identifying Differentially Expressed Genes

Prior to subjecting the data to different comparison statistical to find differentially expressed genes, we wanted to find a way to carry out the study to find dynamically differentially expressed genes. In most studies, the end point of differentiation protocols, HLCs, are compared to primary hepatocytes but due to the complexity of development we wanted to take a holistic approach to look at things dynamically. Our approach was to look at differentially expressed genes by comparing in-vitro differentiation to the course of in-vivo development. Thus, genes with static expression profiles in the dataset were removed. Only those genes changing greater than fourfold across any timewise comparison were

retained in the global analysis. All analysis was carried out in the statistical software, R. To identify orthologous genes whose time dynamic profile is different between hESC differentiation and mouse embryo development, the expression profile of each gene was expressed as a second order polynomial function of the arc that is formed after plotting the mouse and human dataset on a common time scale. The difference of the dynamic expression profiles of a pair of orthologous genes was computed in three ways: Pearson's coefficient, Spearman's coefficient and Euclidean distance.

Further details regarding each of these statistical criteria to identify differentially expressed genes in a dynamic manner are described in detailed below.

The first criteria used for identifying differentially expressed genes was to determine the Euclidean distance between different datasets. The formula used is described below:

$$\boldsymbol{d = \left|(x_{i_m} - x_{i_h}\right| = \sqrt{\sum_{i=1}^{n} |x_{i_m} - x_{i_h}|^2}}$$

Euclidean Distance where Xij represent expression intensities for gene i and species j

The second criteria imposed on the different genes was to use Pearson's correlation described below:

$r = \frac{\sum(x_{i_m} - \overline{x_{\imath m}})(x_{i_h} - \overline{x_{\imath_h}})}{\sqrt{\sum(x_{i_m} - \overline{x_{\imath m}})^2 \sum(x_{i_h} - \overline{x_{\imath_h}})^2}}$ x_{i_h} is the gene in human and x_{i_m} is the corresponding gene in mouse

Pearson's correlation where Xij represent expression intensities for gene i and species j

The third criteria imposed on the different genes was Spearman's' correlation which is described below:

$\boldsymbol{r} = \frac{\sum(y_{i_m} - \overline{y_{\iota_m}})(y_{i_h} - \overline{y_{\iota_h}})}{\sqrt{\sum(y_{i_m} - \overline{y_{\iota_m}})^2 \sum(y_{i_h} - \overline{y_{\iota_h}})^2}}$ y_{i_h} is the gene in human and y is the corresponding gene in mouse

Spearman's correlation where Yij represent the ranks of gene expression intensities for gene i in species j

Genes for which the Pearson's or Spearman's coefficient is ≤ -0.8 were identified as candidates whose expression dynamics between human *in vitro* differentiation and mouse embryo development follow opposite trends; while genes for which the Euclidean distance is > μ+2σ (where μ and σ are average value and standard deviation of Euclidean distance, respectively), were identified as candidates of dynamically differentially expressed genes. Genes with expression values changing fourfold or higher across any sample pair in either human (D5-D20) and mouse data (E9.5-15.5) were further shortlisted as genes represented as highly dynamically expressed different genes.

5.2.4 Different Methods of Transcriptome Analysis

Global analysis to elucidate trends in the data was performed using hierarchical clustering, principal component analysis and non-negative matrix factorization.

Hierarchical clustering analysis was performed on the batch corrected datasets of human and mouse separately using the statistical software R and its *hclust* function. Euclidean distance between the expression value of all gene (*i*) was used as the metric for the distance between different pairs of samples (*a* and *b*).

$$||a\text{-}b||_2 = \sqrt{\sum_i (a_i - b_i)^2}$$

Unweighted Pair Group Method with Arithmetic Mean (UPGMA) was chosen as the distance metric.

$\frac{1}{|A||B|} \sum_{a \in A} \sum_{b \in B} d(a, b)$ where d is the metric chosen (i.e. Euclidean distance)

Another method of whole transcriptome analysis similar to hierarchical clustering is to use principal component analysis (PCA). This is a multivariate statistical technique commonly used to reduce the dimensionality of a dataset, which was performed using the *prcomp()* function in R. Principal component calculation is based on a singular value decomposition (SVD) of the mean centered gene expression data (136). The transcriptome data of all sample was organized in a *n* gene x *m* sample matrix. Let x_{ij} be the expression value of gene i in sample j and A be the n x m mean centered expression matrix, where n is the number of genes and m is the number of samples. The elements of A are thus $a_{ij} = x_{ij} - \overline{x_i}$

By single value decomposition, $A = UEV^T$

Where U is the (n x m) eigenvector matrix of A^TA, E is an (m x m) diagonal matrix and V is the (m x m) eigenvector matrix of AA^T. Elements of U are referred to as eigen vectors and are ordered by their corresponding eigenvalues, which capture the variance of that element. Each eigenvector corresponds to a principal component (PC). The projection of the samples in the principal component space is given by

$$U^TA = EV^T$$

EV^T contains the coordinates or scores for the samples in the PC space, which were calculated using the *prcomp()* function in R. If the first few PCs can capture most of the variance, one can visualize the coordinates of the samples defined by the first few PCs without losing much information. The n x m dimensions of A is therefore reduced to an m x v matrix, where v represents the number of PCs reflecting the largest variance in the dataset. In our analysis, the first two PCs captured about 90% of the variance in all datasets. Therefore, the multidimensional transcriptome data was reduced to m x 2 matrices, where m represents the number of samples in the dataset. The x and y coordinates of this matrix are represented by PC1 and PC2, the components which retain the most variance in the data. These co-ordinates were then used to visualize the ordering of the samples in the two-dimensional PC space.

Another common method for examining transcriptome samples based on its gene expression profile is non-negative matrix factorization (NMF). NMF is another dimension

reduction analysis method (137) used to analyze the batch corrected human or mouse datasets using a NMF R/Bioconductor package (138) . Briefly, *A*, an *n* x *m* transcriptome expression data matrix of *n* genes of *m* samples was resolved as a product of two matrices containing only non-negative values *W* and *H* of size *n* x *k* and *k* x *m* respectively. *k* is the optimal factorization rank or the number of clusters that A can be represented by. This is a critical parameter whose value can usually be determined by plotting the cophenetic coefficient corresponding to different values of ranks, and selecting the rank at which the cophenetic coefficient begins to decrease (137). The samples are then assigned to the number of clusters determined by the cophenetic coefficient by performing several iterations. In the human and mouse NMF analysis, 100 iterations were used to confirm sample assignments.

The elements of matrix W are referred to as metagenes while the matrix H corresponds to the metagene expression profiles (138). To evaluate metagenes that contribute most significantly to the biological variance in the phenotypes of interest, they were assigned scores based on a scoring schema outlined in (139) between values 0 and 1. Higher values correspond to larger contributions of the gene to that cluster. A threshold criteria was set and genes above this threshold were designated as significant metagenes.

5.3 Results

5.3.1 Compilation of Human Differentiation of Stem Cells

The transcriptome data involved in this study is divided into the human dataset and mouse dataset. This meta-analysis consisted of over 15 different studies and 100 samples for the comparison of *in-vitro* differentiation and *in-vivo* development. We carried out a differentiation protocol similar to the one described in the previous chapter on our H9 cell line. The H9 cells were differentiated towards the hepatic lineage using a four-stage hepatic differentiation protocol designed to mimic embryonic development (Figure 5-1A). The transcript of hepatic markers, HNF4A, AFP, ALB, CYP3A4, AAT, and TAT, all increased significantly by the end of the differentiation (Figure 5-1B). Immunohistochemistry reveals that a majority of cells were positive for endodermal marker, FOXA2, by Day 6. Hepatic markers, HNF4A, AAT, AFP and ALB were prominent in cells by Day 20 (Figure 5-1C). The high expression of AFP, which is low in primary hepatocytes, suggests the fetal liver nature of hHLCs derived from PSC. These samples were collected over the course of the

differentiation on day 6, 10, 14, and 20 (D6_H, D10_H, D14_H, D20_H, respectively) and will form the basis for our study and how we can find ways to improve current in-vitro differentiation efforts.

The transcriptome data, along with another study of hHLC differentiation (D9_HE, D12_HE, HLC_H2), were augmented by data from the public domain to increase the data diversity as described more in detail in the previous section. One other important dataset was used as a dynamic comparison because it consisted of multiple time points over their course of differentiation from hPSC to hHLC (Duncan). While other samples are consisted of end point differentiation samples, hHLCs samples, obtained using various differentiation methods. A total of 17,683 genes were common among all the datasets based on ENSEMBL ID.

5.3.2 Batch Correction of Human In-vitro Data

Batch correction was performed on the hHLC *in vitro* dataset to remove the systematic bias caused by the heterogeneous samples and platform sources. The effect of batch correction is visible after performing hierarchical clustering of the different samples before and after batch correction. Without batch correction, samples obtained from each study clustered together, as seen in Figure 5-2. After batch correction, samples based on their similarity in differentiation stages clustered together regardless of the study it came from, suggesting the removal of the batch effects (Figure 5-3A). Samples from a similar differentiation stage from each of the time series studies (Hu/Duncan) clustered together, indicating a similar progression through differentiation. However, the hHLC data from all sources consistently clustered separately from tissue samples of fetal and adult liver samples, confirming the immaturity of these cells.

Non-negative matrix factorization (NMF) analysis was performed to separate the hHLC dataset into three clusters (Figure 5-3B). The optimal number of clusters, 3, was determined by the highest cophenetic coefficient obtained (Figure 5-4). The smallest cluster consisted of mature hepatic samples from liver tissue and primary hepatocytes, and the other two clusters consisted of samples of early (D4-10) and late differentiation (D14-20) stages. Only two hHLC samples (90, 102) were grouped separately in the mature cluster. The grouping of early and late differentiation samples was remarkably similar

between NMF and hierarchical clustering. Notably, samples in similar differentiation stages from the two differentiation time series studies (Hu/Duncan) were clustered together supporting their similar progression through the differentiation stages.

We used Database for Annotation, Visualization, and Integrated Discovery (DAVID) to probe the functional class of the 967 metagenes obtained from the NMF analysis (Appendix Table 9-1) (140, 141). Genes associated with the clustering of early differentiating cell states (D4-10) were involved mainly in developmental processes, while those associated with late differentiation states (D14-20) were mainly involved in cell differentiation, adhesion, extracellular matrix reorganization, epithelial specification and drug response. The metagenes classifying fetal and adult liver tissue were involved in mature liver functions including CYP450 drug detoxification, electron carrier activity, and carbohydrate metabolism, among many others.

Thus, the two unsupervised classification methods, hierarchical clustering and NMF, classified the batch corrected human transcriptome data into similar groups based on the functional relevance of gene expression patterns. This gives credence to the data processing method we adopted for batch correction to further carry out the meta-analysis.

5.3.3 Alignment of Human *in vitro* Data along a Differentiation Scale

The batch corrected transcriptome data of the human dataset was subjected to principal component analysis (PCA) to reduce the number of dimensions of the data. Two principal components, PC1 and PC2, captured 90% of the data variance, suggesting that two components were sufficient to display the variability within the transcriptome data of the samples. PC1 and PC2 of each sample were then plotted on a PC1 vs. PC2 graph (Figure 5-3C). All samples, starting from endodermal cells to HLCs were observed to chronologically align along an arc while primary hepatocytes and adult liver were located in the high PC2 region. The two time series data sets (Hu/Duncan) both aligned from left to right in chronological order based on the differentiation stages. Interestingly, the iHEP and the liver bud samples aligned with earlier differentiation samples (D10-14) of the time series analyses. All HLC samples, regardless of the laboratory and protocol used, aligned within a narrow region in the principal component space, suggesting that they all have a similar degree of hepatocyte maturity.

5.3.4 Compilation and Alignment of Mouse Fetal Liver Development

The data processing pipeline described for the human datasets was also used to process the dataset of mouse embryonic liver development. The batch corrected data from two studies composed of 47 samples across the development of two strains of mice during embryo development. After compilation, a total of 16,415 genes were common among all the samples. These samples were then subjected to hierarchical clustering, NMF and PCA as we had done before for the human *in-vitro* dataset. Similar to the observation made before, batch correction removed the bias arising from different platforms and sources (Figure 5-5,5-6). Clustering of the samples into two clusters gave the most optimal grouping compared to three or four clusters, as assessed by the cophenetic coefficient (Figure 5-7). NMF classified the batch corrected dataset into two groups, early and late development, which was in line with the hierarchical clustering results (Figure 5-6). When the data was subjected to PCA, the first two principal components captured over 90% of the data variance. As was seen for the hPSC-HLC time series, we found that the mouse embryonic liver development samples were lined up chronologically in order of their developmental stage (Figure 5-6).

From the NMF and PCA analysis, the samples seemed to be grouped into an early (E9.5-E14.5) cluster and a late (E15.5-E19.5) cluster. This classification agrees with our understanding of liver development, where the E14.5 stage represents a transition from primarily hematopoiesis to hepatocyte maturation. Functional analysis on the 129 metagenes obtained from the NMF grouping confirmed that genes related to liver functions were key in the grouping of samples (Appendix Table 9-2).

5.3.5 Integration of Mouse *in vitro* Differentiation and *in vivo* Development Data

To determine whether or not it was possible to combine *in-vitro* data with *in-vivo* data for studying the dynamic of gene expression, we wanted to first combine datasets originating from the same species. Thus, transcriptome data of miPSCs differentiating to HLCs at different timepoints, as well as murine fibroblasts being reprogrammed to hepatocytes (miHeps), were integrated with the mouse embryonic development data sets by performing batch correction treating the *in vivo* and *in vitro* samples as two separate batches. Again, batch correction removed the bias of platforms, sources and *in vitro/in vivo*

contrast among the different samples. These combined data were then subjected to PCA like before. On a space of PC1 vs. PC2, the *in vitro* differentiation data and *in vivo* development data can be seen to line up along an arc based on the maturity of the samples (Figure 5-8). The mouse iPSC differentiation data aligned chronologically with E9.5 to E15.5 stages of mouse liver development, while the data points for mouse embryonic liver development beyond E15 continue to spread to the region with higher values of PC2. The results suggest that the miPSC-HLCs and miHEPS acquired a differentiation stage equivalent of ~E13.5-15.5 stage and still lack the maturity of further developmental stages, let alone fully developed prenatal E19 and postnatal liver samples.

5.3.6 Development of Unified Time Scale Using Mouse *in vivo* and Human *in vitro* Dataset

Having shown that mouse iPSC-derived HLCs and mouse embryonic liver development data can be integrated using batch correction and projected onto a PC1 vs. PC2 plane for comparison), we next integrated the hPSC *in vitro* differentiation data with mouse *in vivo* embryo development data. The batch corrected mouse embryonic liver development (E9.5-19.5) transcriptome data was combined with batch corrected hPSC-HLC differentiation transcriptome data. Common homologous identifiers using the BioMart database package found in the statistical program, R, generated a master dataset with 14,312 genes. Batch correction was performed on the combined dataset using the ComBat function in R to eliminate the batch effects from the biasing of different species between the human *in vitro* and mouse *in vivo* samples.

Comparison of the expression levels of the hepatic markers, alpha-fetoprotein (*AFP*) and albumin (*ALB*), for mouse development and our time series of human *in-vitro* differentiation (Hu) showed consistent trends in gene expression dynamics before and after batch correction (Figure 5-9). After batch correction, the expression levels were slightly shifted down but the dynamics of the expression through the stages were preserved for each species.

PCA analysis was performed on the integrated dataset of h/mPSC-HLCs and mouse embryonic liver development. PC1 and PC2 for each sample from both species was plotted onto a two dimensional PC space. The data points of mouse and human aligned along the

same arc, and lined up in order of their development or HLC differentiation stages respectively (Figure 5-10).

From the alignment of hPSC-HLC differentiation and mouse embryonic liver development data, the corresponding stages of human HLC differentiation and development of mouse liver can be deduced. hPSC derived endodermal cells of differentiation were aligned with E9.5-E10.5 samples, but the majority of fully differentiated HLCs were aligned with E13.5-E15.5 of mouse development. These results strongly indicate that HLCs derived from hPSCs are more similar to the mouse ~E14-E15 stage of *in-vivo* development, similar to what was observed with HLCs derived from mPSCs.

5.3.7 Differentially Expressed Genes in HLCs vs Mature Cells

The transcriptome of all HLCs for human (hHLCs) and mouse (mHLCs) were separately compared to their mature state (primary hepatocytes and E19.5 liver, respectively) using Significance Analysis of Microarray (SAM-R) in R, where the criteria of $q<0.05$ and fold change > 4 were imposed to identify differentially expressed genes. 129 genes were identified as differentially expressed between hHLCs and mature cells (primary human hepatocytes (PHHs), fetal and adult liver). Among the 129 differentially expressed genes, processes involved in CYP450 drug metabolism, carboxylic acid, amine and lipid metabolism, and complement and coagulation cascades were identified as enriched using the DAVID tool (141). A similar differential expression pattern was seen when comparing mHLCs to E19.5. The mHLCs and E19.5 comparison identified 127 genes as significantly differentially expressed, of which 42 genes were common between the human and mouse analyses, which are shown in Figure 5-11 and listed in Appendix Table 9-3. Mature hepatocyte genes differentially expressed in both species included metabolic genes, such as *G6PC* and *FBP1*, and the cytochrome P450 enzymes, *CYP3A4* and *CYP2C9* indicating the lack of the maturity in the metabolic profile of HLCs from both species.

5.3.8 Expression profile comparison on a Unified Developmental Time scale

The alignment of human *in vitro* HLCs and mouse *in vivo* embryonic development data on a common coordinate grid presents the unique opportunity to compare the dynamics of

their developmental events along a common “time” scale. We treated the arc formed by those data points as a common developmental path. Using the first point (E9.5) as a reference point, the distance of a sample from the reference point along the developmental path can be taken as a “unified developmental time (DT)”. Using the equation shown below, where x corresponds to PC2 values of the samples for that dataset, DT can be computed for each sample, irrespective of its sample origin.

$$P(x) \simeq a + bx + cx^2$$

and now compute the arc length

$$\int \sqrt{1 + (P'(x))^2}\, dx = \int \sqrt{1 + (b + 2cx)^2}\, dx$$

$$\int \sqrt{1 + (P'(x))^2}\, dx = \frac{(b + 2cx)\sqrt{(b + 2cx)^2 + 1} + \sinh^{-1}(b + 2cx)}{4c}$$

where x corresponds to PC2, P(x) corresponds to the polynomial fit of the curve in the PC1/PC2 space and a ,b and c correspond to the coefficients of the polynomial function.

A higher value of DT indicates a higher maturity for that sample. The PC2 component of all the human samples along with the mouse developmental samples were plotted against their DT as shown in Figure 5-12. The two human HLC time series datasets (Hu/Duncan) both showed a progressive movement along the DT to approximately 0.4 (equivalent to E15.5) by the end of the differentiation.

To understand the dynamics of certain genes during development and *in-vitro* differentiation, we combined the samples of the two human time series PSC-HLC differentiation (Hu/Duncan) into one human in-vitro process and plotted it along with the mouse development data on the same DT scale. While some hepatocyte genes (e.g. *AFP*) showed similar trends between human *in vitro* and mouse *in vivo* data, others showed inconsistent trends. Dynamic differential expression analysis using Pearson correlation, Spearman’s coefficient and Euclidian distance (PSE) was performed to compare the dynamic between mouse development with *in-vitro* differentiation for each of the 14,332 orthologous genes. The Pearson correlation coefficient measures the linear correlation, while the Spearman’s coefficient measures the monotonic relation between two variables. Briefly, the threshold criteria were set for Pearson and Spearman coefficient $\leq$ -0.8 and

Euclidean distance $\geq \mu+2\sigma$. The genes fulfilling the PSE criteria were plotted and visually inspected for both human *in-vitro* differentiation and mouse development (Figure 5-13).

5.3.9 Functional Analysis and Comparative Gene Expression

197 genes were found to be dynamically differentially expressed between PSC-HLC differentiation and mouse embryonic liver development using the PSE criteria (Appendix Table 8-4), including 11 transcription factors (TFs) (Figure 5-12), 7 transporters and 33 other cell surface markers. The major functional classes/ pathways involved from the 197 genes identified by the DAVID tool, included developmental process, organ development, cell adhesion as shown in Figure 5-14.

Many transcription factors regulating organ development, including *HAND1, PITX2, ALX1, CDX2, TSHZ1* and *SNAI2*, were highly expressed in the early stages of mouse liver development and subsided after E15. However, their transcript levels in during PSC differentiation followed an opposite trend and remained high in HLCs (Figure 5-12). To confirm the dynamic of the differentially expressed TFs, we also measured the transcript level over the course of differentiation using qRT-PCR following the Hu differentiation. (Figure 5-15). *ALX1* is involved in forebrain development (142), *CDX2* plays a role in lineage segregation of the inner cell mass and trophectoderm (142, 143) as well as intestinal fate and epithelial mesenchymal signaling (144). *HAND1* regulates embryonic cardiac development (145), *PITX2* participates in limb development (146, 147), *SOX9* in pancreatic and bile duct development (148), *SNAI2* in epithelial mesenchymal transition (149, 150) and *TSHZ1* in pancreatic cell development and maturation (151). Furthermore, TGFβ signaling in conjunction with upregulation of *SNAI2* are consistent with epithelial to mesenchymal transition (EMT), which has also been described purely as being abnormally expressed in HLCs compared with PHHs (149). The opposing trend of transcript dynamics of these TFs may be indicative of discordant gene regulation between mouse liver development and HLC differentiation. We also examined the TF binding sites in the promoter regions of the dynamically differentially expressed genes (http://genome.ucsc.edu/ENCODE/). The TFs that have their binding site present in the promoter region of a large number of differentially expressed genes are listed in Table 5-3.

Several cell surface proteins also had opposite trends of transcript dynamics between HLC differentiation and mouse liver development including *SPARC*, an extracellular matrix secreted factor, *CDH3* (*P-Cadherin*) and the cell surface receptors, Frizzled (*FRZB*) and *TGFBR3* (for WNT and TGFβ signaling respectively). Their transcript levels increased during the course of differentiation, but decreased in the developing embryonic mouse liver by E15.5.

Expression of *GLUT2* (*SLC2A2)*, a liver specific glucose transporter, increased in the developing embryonic mouse liver but not in HLCs. The transcript dynamics of *GLUT2* was plotted together with that of mature human cell types, and mouse *in vivo* liver development on the unified developmental time (Figure 5-16). While *GLUT2* transcript levels gradually increased in mouse liver development, it remained relatively low in differentiating HLC. A similar contrasting expression pattern between mouse liver development and HLC differentiation was also seen for three other enzymes involved in glucose metabolism that are highly expressed in liver: aldolase B (*AldoB)*, glucose-6-phosphatase (*G6PC)* and fructose-1,6-bisphosphatase (*Fbp1)*. The latter two are involved in gluconeogenesis. Using the pipeline we established to compare *in vivo* developmental data with *in vitro* differentiation data, we were able to clearly show that the identified differentially expressed enzymes that are not expressed in HLCs are characteristic of mature hepatocyte.

5.4 Discussion

In this study, we used highly heterogeneous transcriptome data from human and mouse PSC *in vitro* differentiation and mouse *in vivo* embryo development for meta-analysis. Due to the heterogeneous nature of the dataset, batch correction needed to be performed to remove the batch effects. Previous studies have shown that the empirical Bayes based ComBat algorithm can be relied on in removing systems bias caused by a mixture of transcriptome data from different populations of cells of different sources (152-157). Through successive data processing with increasing level of data complexity, we successfully integrated hPSC-HLC differentiation data from different studies with mouse *in vivo* liver development data. In each level of analysis, after batch correction, unsupervised clustering grouped the data by their similarity in the stage of *in vitro*

differentiation or by their embryo developmental day. The batch-correction algorithm employed was originally developed to deal with the effects of different sources and platforms. We extended its use to combine data of different species. The clustering results of the batch corrected data give credence to the validity of the data processing. We examined the identity of the metagenes used in NMF and found them to be related to development and liver functions, confirming the biological relevance of clustering of the samples.

Two observations can be made from the PCA results: first, differentiating HLCs align along their differentiation progression in a PC1 vs. PC2 plot; second, the HLCs derived using different protocols and from various sources, for both human and mouse, all lined up in a region that is composed of differentiated, but still immature fetal liver cells of E13-15. This indicates that irrespective of the species of origin, or the protocols used, differentiating HLCs appear to encounter universal roadblocks preventing their maturation. It also suggests that there might be a fundamental block that prevents HLCs from maturing further.

Because all the time series data points of both *in-vitro* and *in-vivo* samples align along a common trajectory in the PC space, the results also suggest that the position along the trajectory can be a common measurement of the developmental stage for *in vitro* and *in vivo* samples from both species. We called this measurement “unified developmental time”.

Identification of similar stages of development across different species that have vastly different time scales often relies on common morphological, biochemical and genetic hallmarks. For example, in human embryo development day 22 and 52 are considered to be equivalent to mouse E9.5 and E14.5 respectively (158). In an earlier study, neural development events were codified and used to generate a regression model for predicting a translation table of times of corresponding stages across nine species(159). Recently, transcriptome data of nematode species were compared to morphological markers to establish embryo developmental milestones in different species of *Caenorhabditis* (160). In another study, different feature measurements of leaves for different tomato species were subjected to principal component analysis to establish a

developmental trajectory (161). Similarly, an algorithm taking RNA-seq data of a species' developing embryo and subjecting them to principal component analysis for alignment in the PC plane was reported to describe the developmental stage (162). However, in many cases, the maturity of HLCs is measured by comparing it to primary hepatocytes at the functional, proteome, and transcriptome level. To our knowledge, this is the first study to compare genes across species and across the *in vitro/in vivo* border to identify the developmental stage at which HLCs are stuck at.

Our meta-analysis revealed that all HLCs, regardless of the differentiation protocol used or species, were closer to mid-liver development (~E14.5) than to primary adult hepatocytes. We aligned the transcript dynamics of human *in-vitro* PSC-HLC differentiation and mouse *in-vivo* development on the "unified developmental time" to identify genes whose expression was progressing in opposite directions. Among the genes that were differentially expressed between h/mHLCs and mature cells (E19.5, adult liver) were key *CYP450* enzymes (*CYP2C8, CYP2C9, CYP2E1, CYP2D6* and *CYP3A4*) and metabolic isozymes (*ALDOB, GLUT2, G6PC* and *FBP1*) that are all hallmarks of mature liver metabolism.

The dynamically differentially expressed genes between hHLC differentiation and mouse liver embryo development encompass many hallmarks of the lack of maturity in HLC. An example is the expression of genes involved in the glucose metabolism. As hepatocytes become more mature, a new set of transporters, enzymes and their isoforms in glucose metabolism are expressed to give the liver its capability in maintaining the homeostasis of glucose and gluconeogenesis. Notably, the increased expression of glucose transporter *GLUT2*, and aldolase B (*ALDOB*) seen in mouse liver development and human adult hepatocytes did not yet occur in HLCs (Fig S11). Not surprisingly, the increased expression of the gluconeogenesis enzymes, glucose 6-phosphatase (*G6PC*), fructose 1,6-biphosphatase (*FBP*) seen in mouse development was not seen in HLCs. Furthermore, we observed that 100 of the dynamically differentially expressed genes were found to have the same dynamic in both mouse and human in-vitro differentiation (Appendix Table 8-5). This observation informs us that although not all genes follow the same dynamic between

mouse and human differentiation, this does give us an insight into the key genes that need to be focused on and understanding how it is contributing to the immaturity in our HLCs.

The gene expression profile and metabolic activities has led to the notion that HLCs were closer to fetal than adult hepatocytes (124). Another study of comparison between HLCs and human hepatocytes at the transcriptome level revealed 4000 differentially expressed genes (123). Our study demonstrates for the first time that HLCs obtained using various protocols and from different sources are all equivalent to E13-15.5 of mouse development. We were able to pin point the stage of developmental block by developing a "unified developmental time" that can span across species and *in vitro* and *in vivo* samples. Despite the difference in the methodology used, a large fraction of genes identified as differently expressed in the previous studies (123, 124) were also identified as dynamically differentially expressed in our study (Figure 5-17). However, differences between the two studies are also seen. Given the different reference of comparison used, (time course of embryonic liver development in this study vs. primary hepatocytes (123, 124), some differences in the identified gene list are not surprising. Most studies conclude the maturity of HLCs by measuring the expression of mature markers. But a more accurate measurement of maturity would be to compare the dynamics to an *in-vivo* development dataset where we are not only able to look into liver markers that are not expressed high enough but also at certain genes that might need to be silenced or are behaving erratically.

Interestingly, a number of common transcription factors identified as differentially expressed are not known to play a role in hepatocyte differentiation, but are involved in the development of heart (*HAND1*), or heart, eye, lung (*PITX1*). Although this may not be surprising, as hepatocytes are created via a mesendodermal precursor early during development, persistent expression of these TFs may prevent final maturation of HLCs to the level of primary hepatocytes.

This study reaffirmed that HLCs are at an immature hepatocyte cell state, and for the first time identifies the "corresponding state" during embryonic liver development. Many genes whose expression followed a different pattern from that in the developing embryo might contribute to this block in further maturation. Whether inhibition of mis-expressed TFs and activation of TFs that are missing will enhance maturation of stem cell

derived hepatocytes remains to be determined. In addition, this meta-analysis pipeline should be applicable to PSC differentiation studies to other lineages if the time course transcriptome data for both in vitro differentiation and in vivo development are available.

5.5 Figures

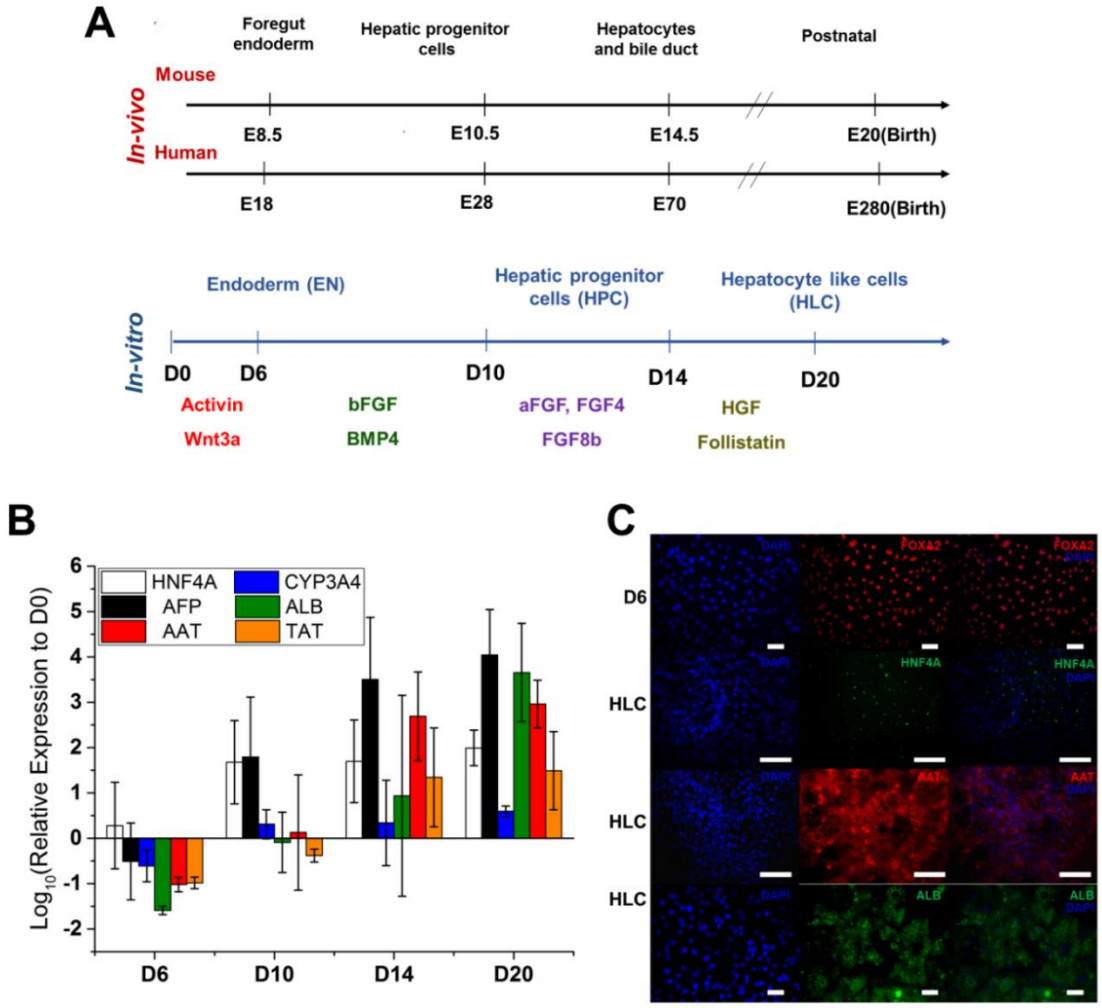


Figure 5-1. Differentiation of human pluripotent stem cells to hepatocyte-like–cells. A) The timelines for mouse and human embryonic liver development and *in- vitro* stem cell differentiations. B) Quantitative PCR results showing the expression levels of key hepatic markers during the course of hepatocyte differentiation depicted in A. C) Immunostaining of the endodermal marker, FOXA2, during the endoderm stage of differentiation and several hepatic markers (Hepatocyte nuclear factor 4-alpha (HNF4A), Alpha-1-Antitrypsin (AAT), Albumin (ALB)) at the last day of differentiation Scale bar. 50um.

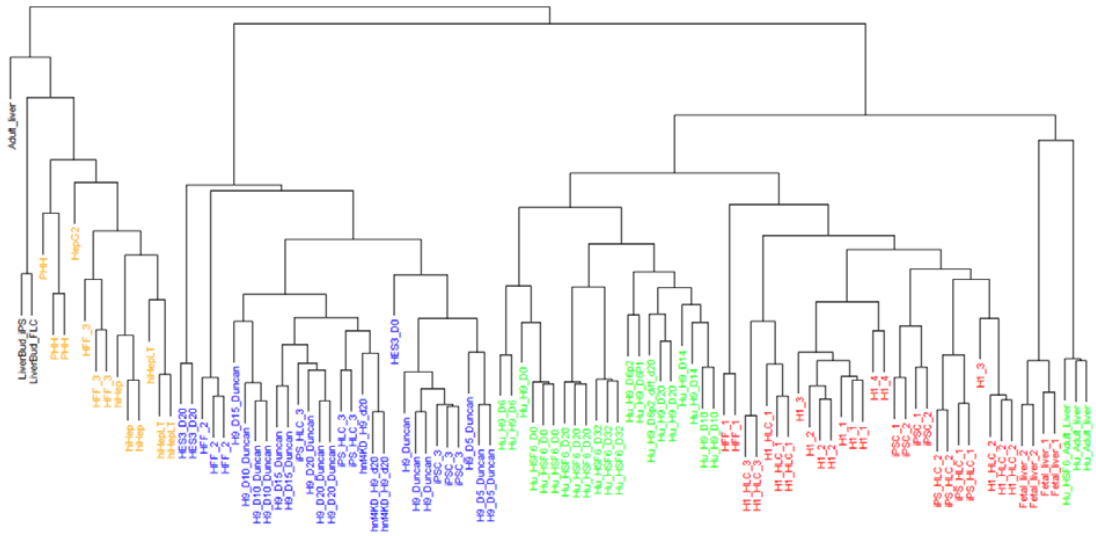


Figure 5-2: Hierarchical clustering of human *in-vitro* differentiation data before batch correction. Platform dependent batch effect were evident for human *in-vitro* differentiation data. The different colors indicate the different platforms for each sample.

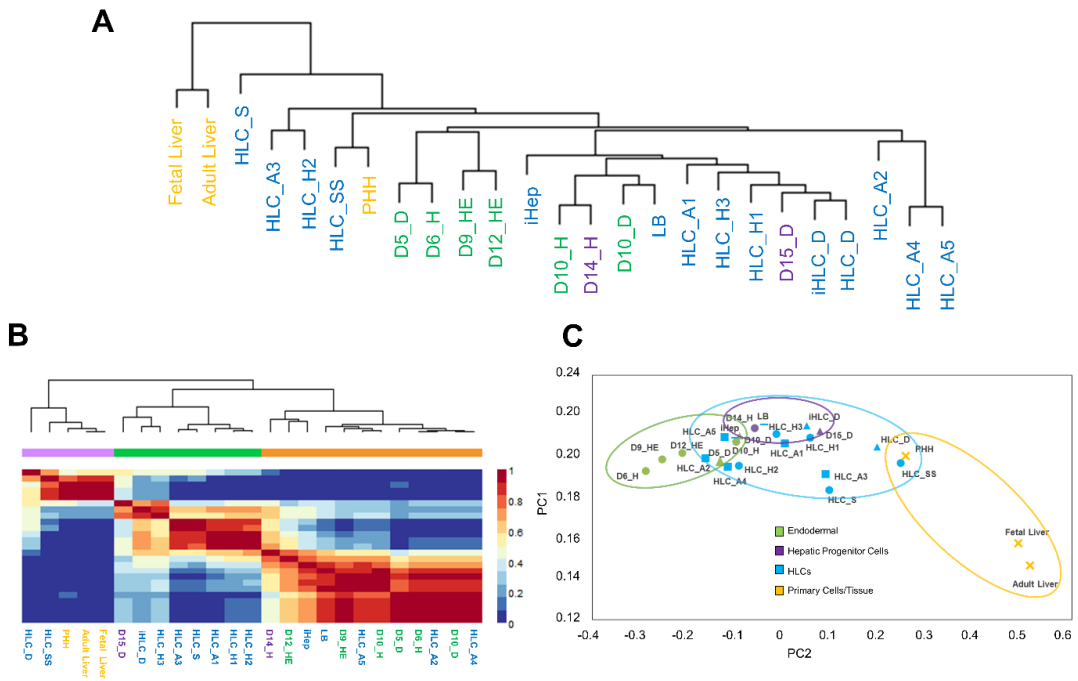


Figure 5-3. Clustering of *In-vitro* Human Differentiation Data After Batch Correction. A) Hierarchical clustering of samples at different stages of differentiation. The colors represent different sources of data. See Table 5-1 for sample annotation. B) Non-negative matrix factorization on human *in-vitro* data after batch correction. All the samples clustered into three distinct groups of early, mid, and late stages of differentiated cells. C) Principal component analysis. The samples aligned according to their differentiation status in the principal component space, starting with endodermal cells followed by HLCs and the most mature cell types to the far right. The colors represent the different stages of the samples and the shapes represent the different sources of the data.

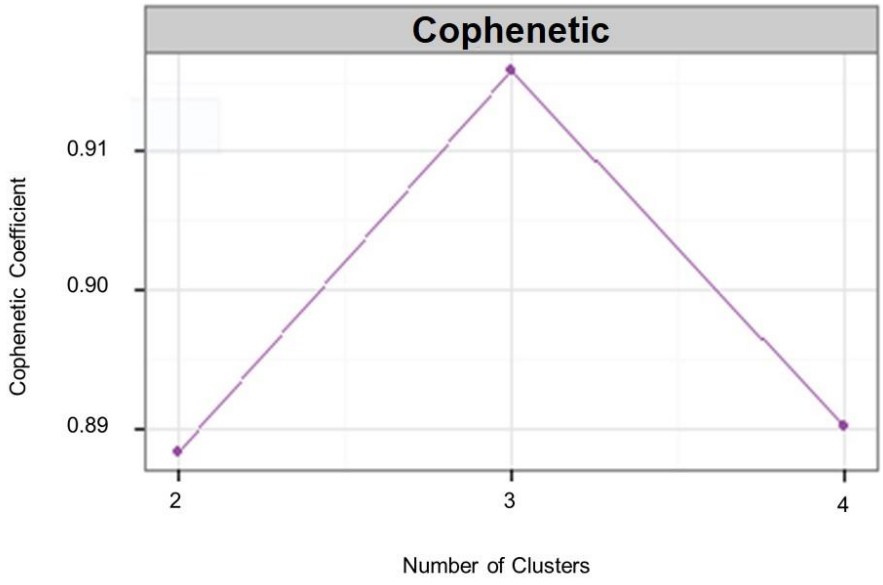


Figure 5-4: Cophenetic plot for human samples. The cophenetic coefficient for the human differentiation dataset is plotted for different ranks. Since the coefficient begins to decrease at 3, the data was optimally clustered into three groups.

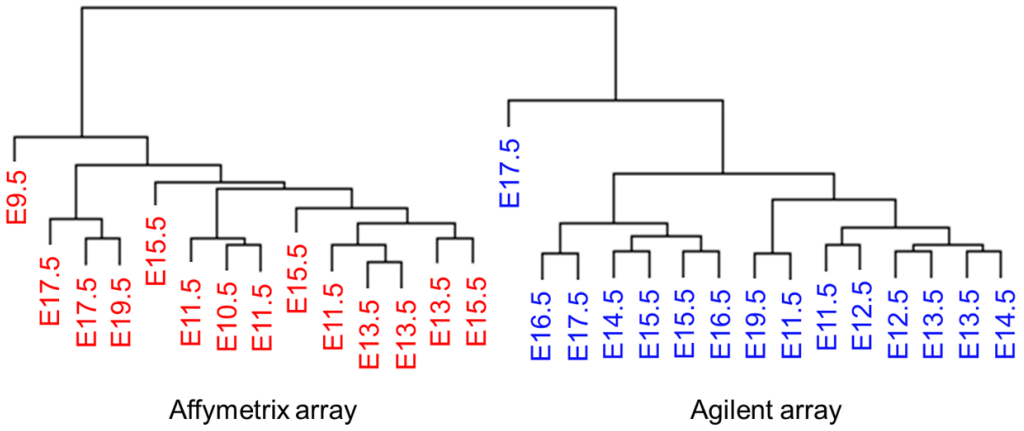


Figure 5-5: Platform dependent effects of mouse developmental data. The two sets of mouse development data where both platforms are represented in red and blue are illustrated. The transcriptome data were clustered largely by their sources and platforms without batch correction.

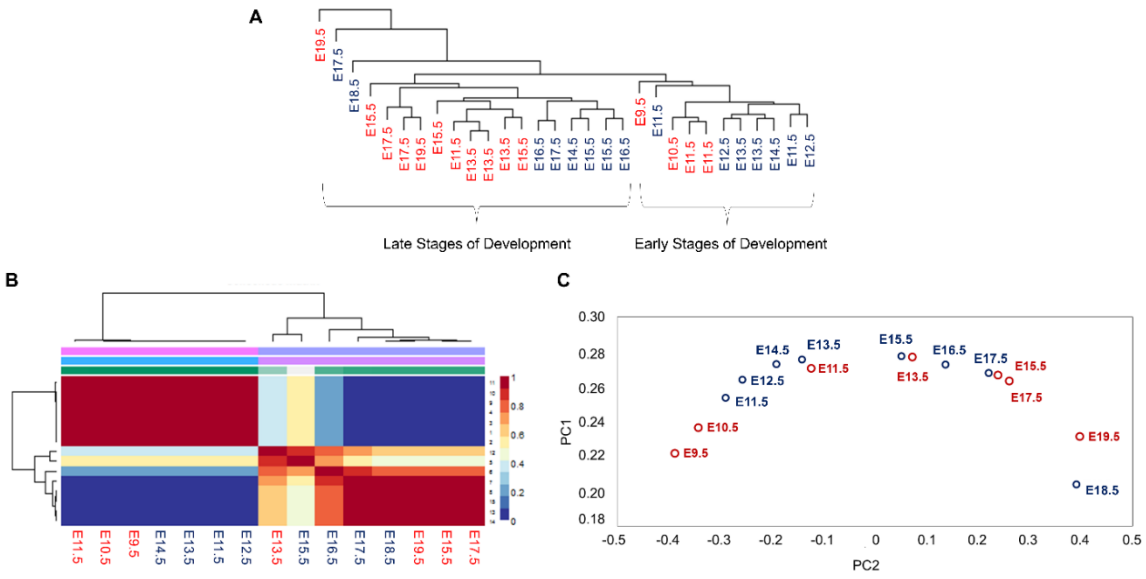


Figure 5-6. Clustering of Mouse Embryonic Development Data After Batch Correction. A) Hierarchical clustering grouped different developmental data into early and late stages of development. (red): isolated $CD34^{-}Ter119^{-}$ cells from fetal liver, (blue): cells from a whole liver . B) Non-negative matrix factorization on mouse embryonic development data (biological replicates from (A) were averaged) separated the mouse data into two groups of early and late stages of development. C) Principal Component analysis aligned samples (biological replicates from (A) were averaged) according to their developmental stage in the principal component space.

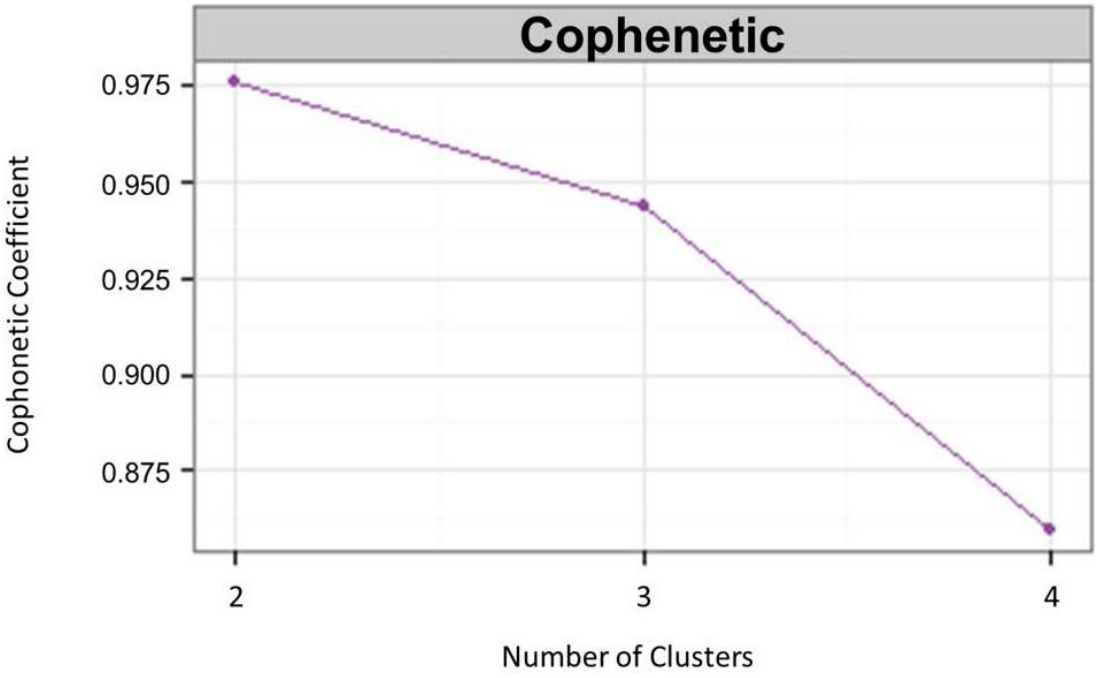


Figure 5-7. Cophenetic Plot for Mouse Development Samples. The cophenetic coefficient begins to decrease at 2 for mouse development, therefore data was optimally clustered into two groups

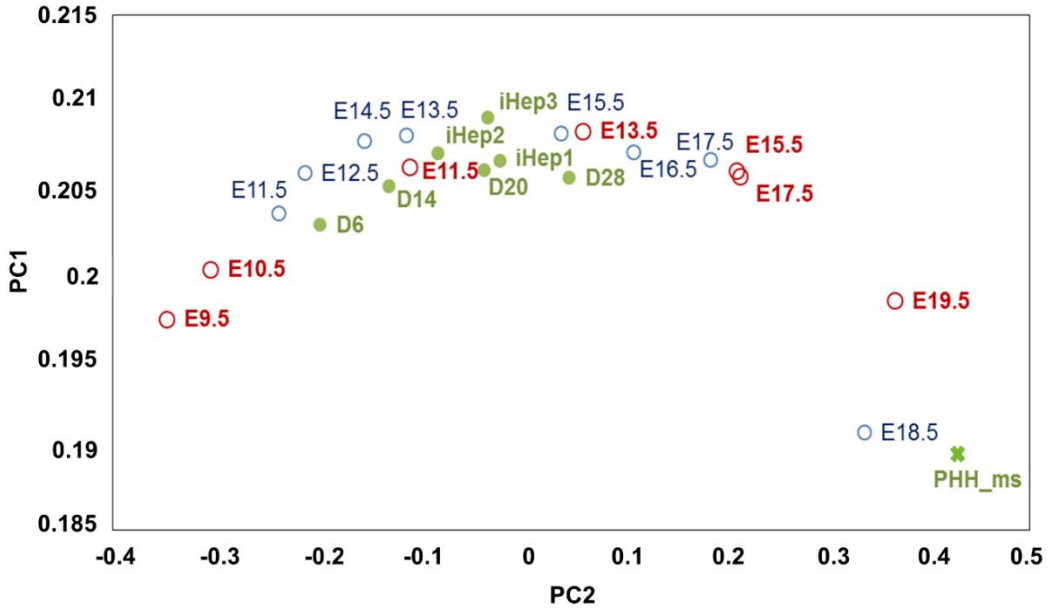


Figure 5-8. Principal Component Analysis on Combined Mouse *In-Vitro D*ifferentiation and Mouse Embryonic Liver Development Data.

Open circles represent embryonic developmental data, blue and red indicate samples from two different studies (Figure 5-6). Closed green circles represent samples from *in vitro* differentiation and X represents primary mouse hepatocytes. To confirm the robustness of our approach, a separate mouse *in-vitro* dataset was integrated with *in-vivo* data for studying the dynamic of gene expression. More specifically, transcriptome data of a time course analysis of miPSCs differentiating to HLCs and murine fibroblasts being reprogrammed to iHeps were integrated with the mouse embryonic development data sets by batch correction. Again, batch correction removed the bias of platforms, sources and *in vitro/in vivo* contrast. The mouse iPSC differentiation data aligned chronologically with E9.5 to E15.5 stages of mouse liver development, while the data points for mouse embryonic liver development beyond E15 continue to spread to the region with higher values of PC2. The results suggest that the miPSC-HLCs and miHEPS acquired a differentiation stage equivalent of ~E13.5-15.5 stage and still lack the maturity of fully developed prenatal E19 and postnatal liver.

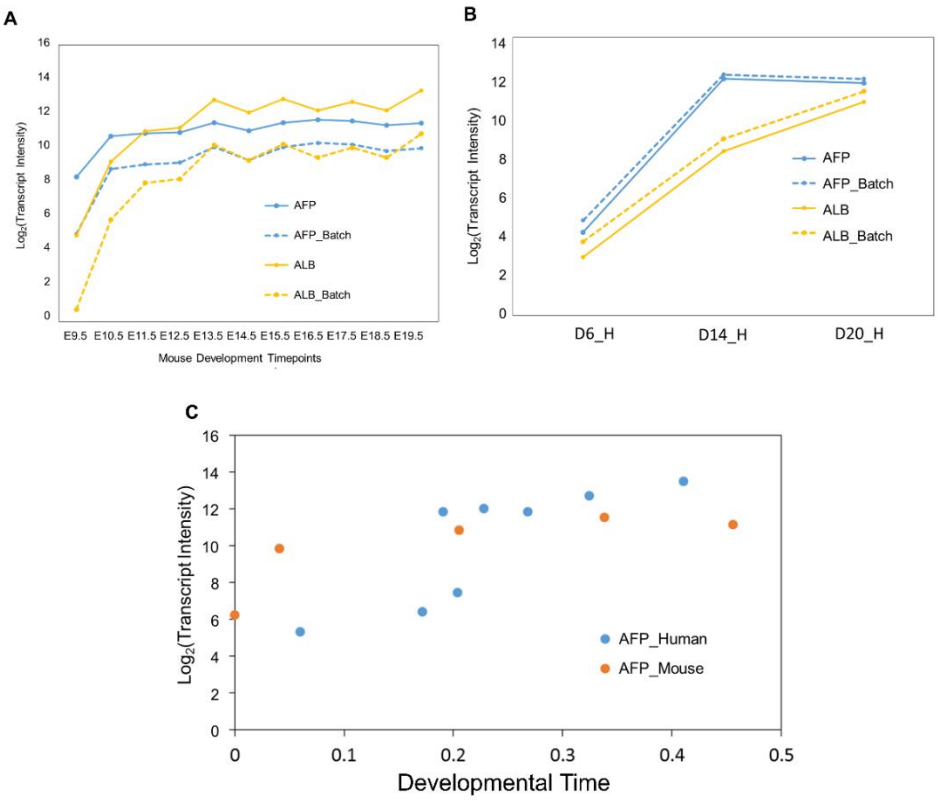


Figure 5-9. Gene expression profile in mouse developmental samples before and after combining the human *in vitro* and mouse *in vivo* data by performing batch correction for species effects. A) The gene profiles for AFP and ALB for mouse samples (solid line) are shifted but the dynamic trend is preserved after batch correction (dash line). B) The gene expression profile in human differentiation samples before (solid line) and after (dash line) combining the human *in vitro* and mouse *in vivo* data by performing batch correction for species effects. C) The gene expression profile for AFP in mouse and human samples after batch corrections

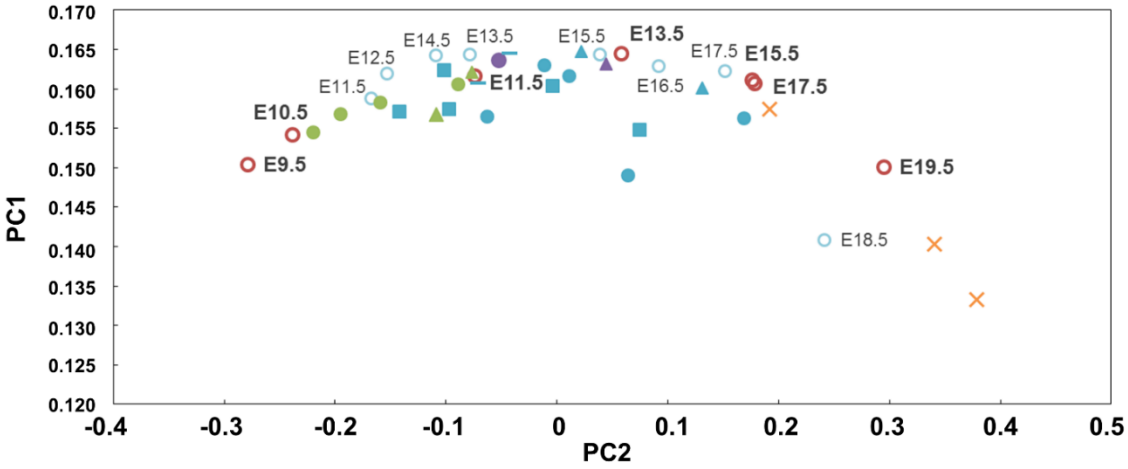


Figure 5-10. Principal Component Analysis on Combined Human *In-Vitro* Differentiation and Mouse Embryonic Liver Development Data. Combined human *in-vitro* and mouse *in-vivo* datasets were subjected to the principal component analysis. The blue and red circles of mouse samples represent two different studies. Human *in vitro* differentiation data are represented by color according to their differentiation state. Endodermal cells, hepatic progenitor cells and HLCs samples are in green, purple, and blue, respectively. The circles, squares and triangles correspond to samples from different studies described in Table 5-1.

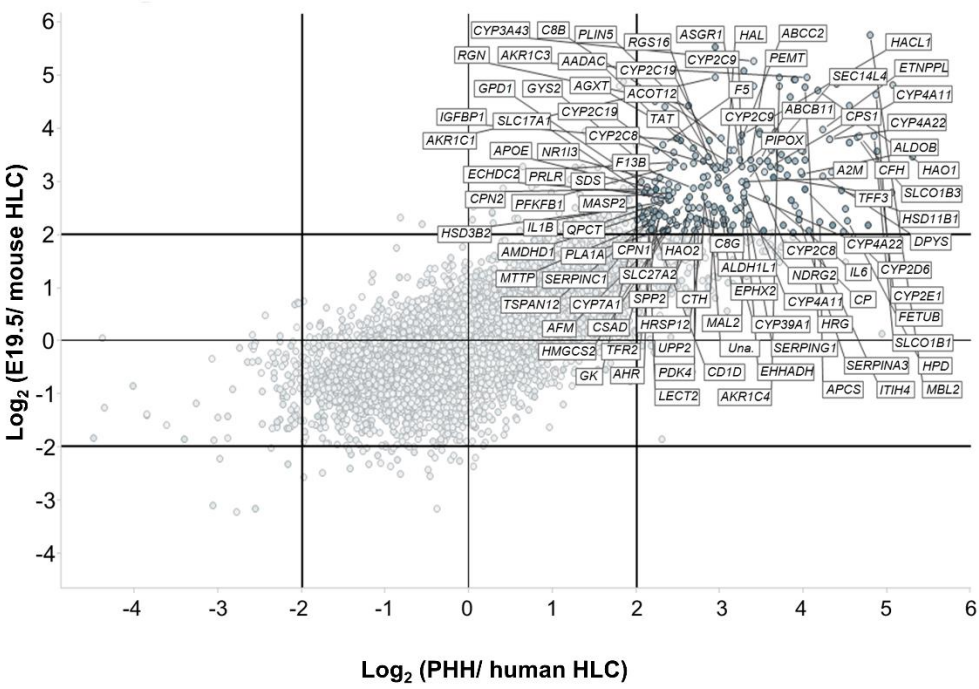


Figure 5-11. Differentially Expressed Genes between HLCs and PHH. Transcript levels of orthologous genes of human and mouse HLC were compared to that in PHH and E19.5 embryo, plotted against each other. Genes consistently differentially expressed (upregulated by four fold or higher in E19.5) in both comparisons were highlighted. *Una.- Unannotated gene.

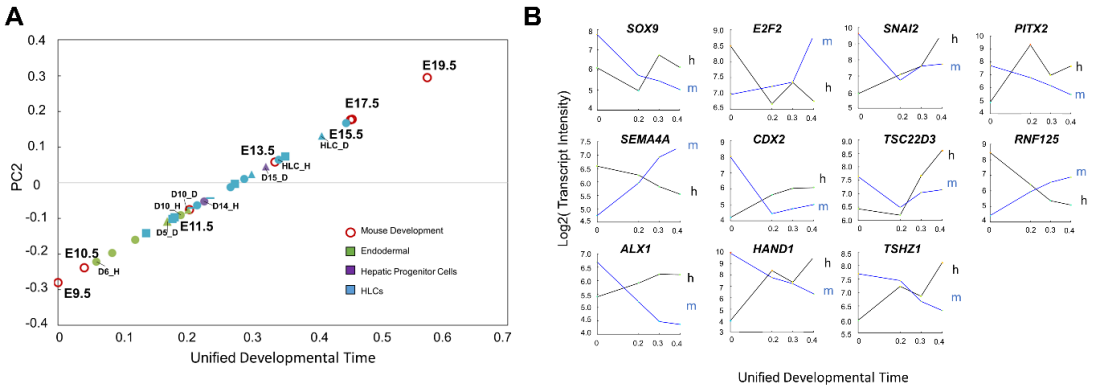
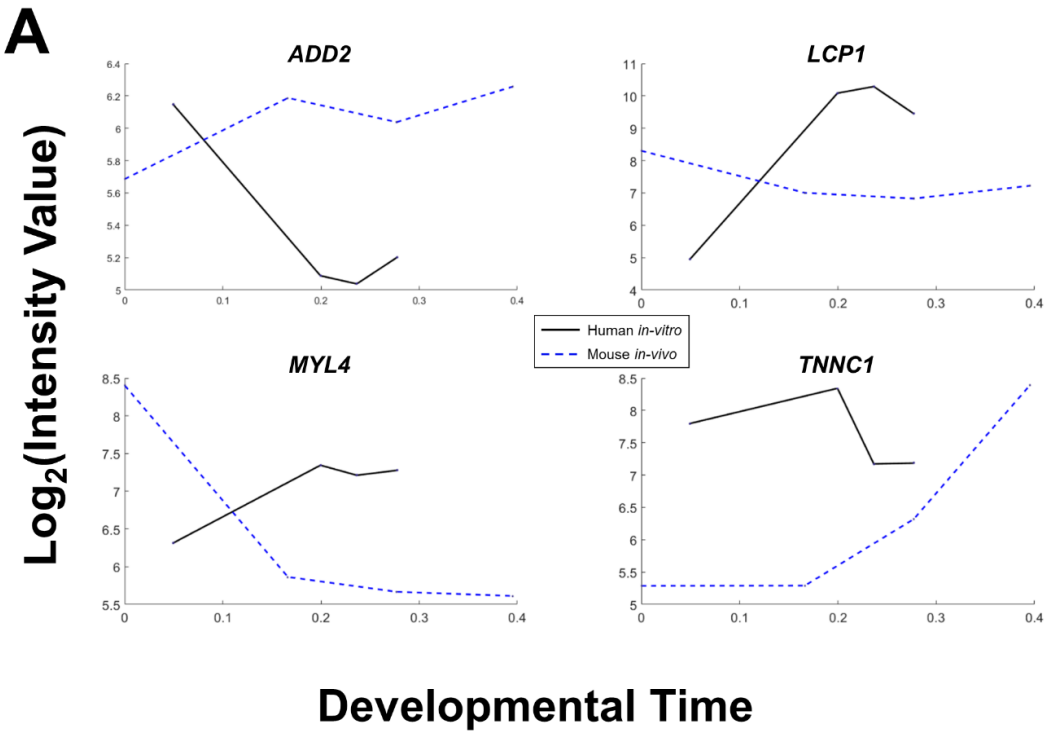
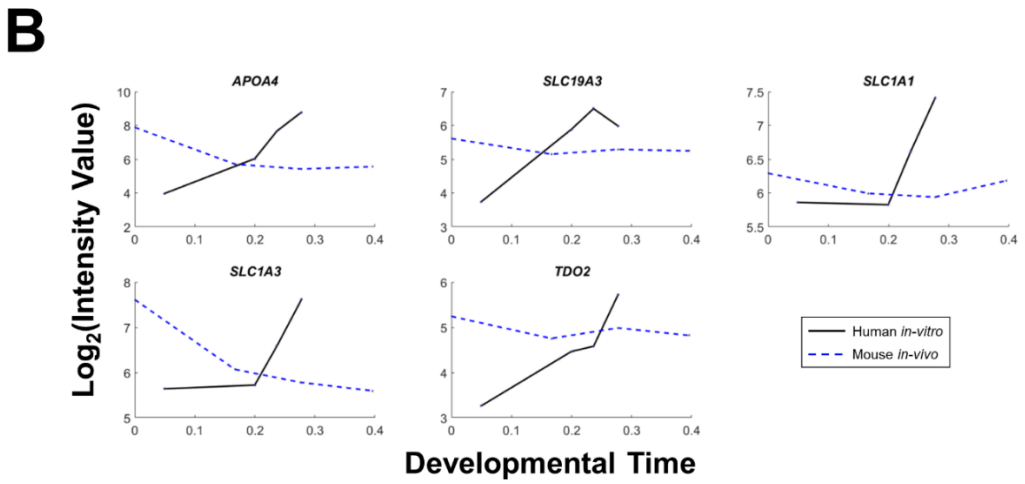
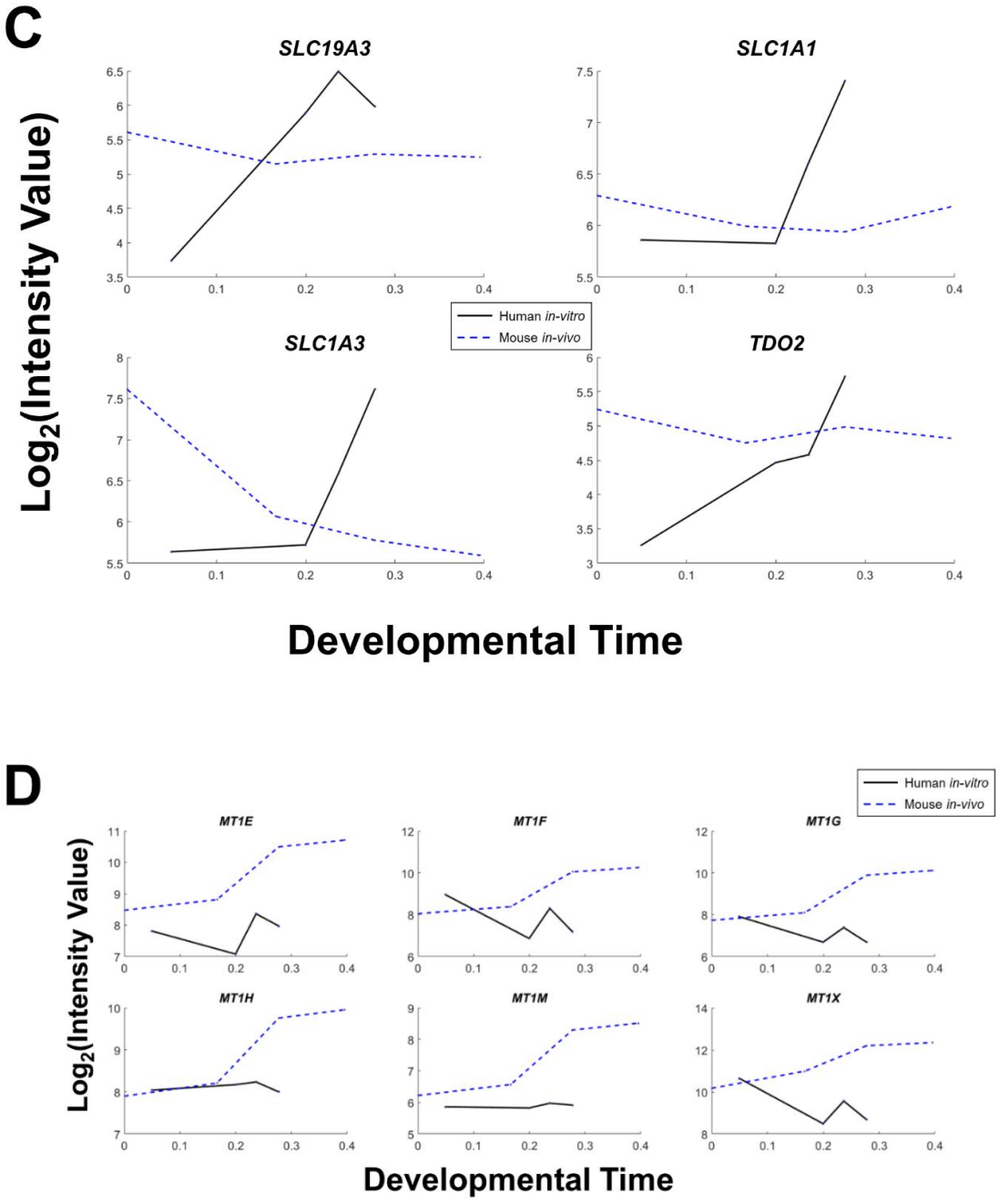
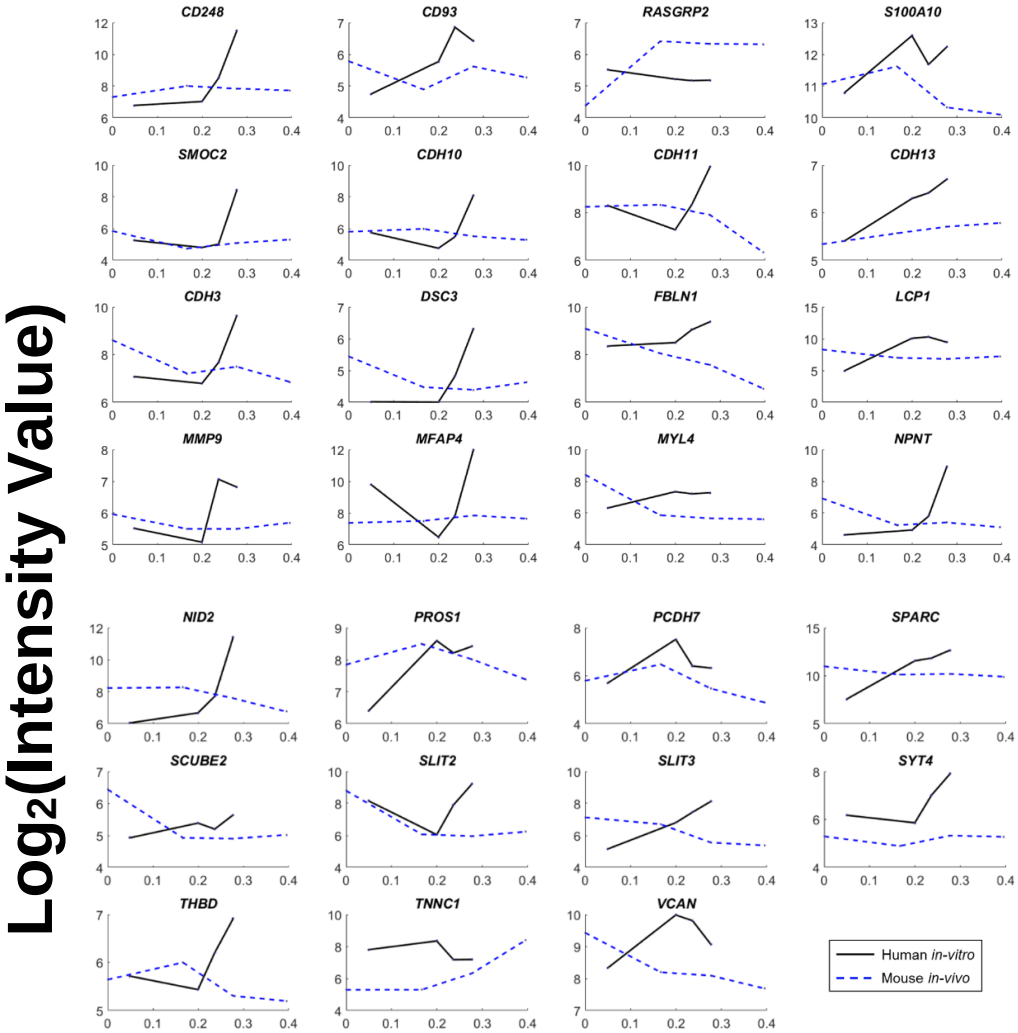


Figure 5-12. Alignment of Samples Along the Developmental Time and Identification of Dynamically Discordant Transcripts. A) Mouse embryo development and human *in vitro* differentiation samples aligned along Unified Developmental Time (DT). The PC2 values for mouse and human samples are plotted against DT. The DT can be represented as a common measure of maturity for *in vitro* differentiation and *in vivo* development. B) Cross-species/Cross-in vitro-in vivo Comparison of transcript dynamics based on DT as a common time scale. Notable Transcription Factors are plotted along the Developmental Time.

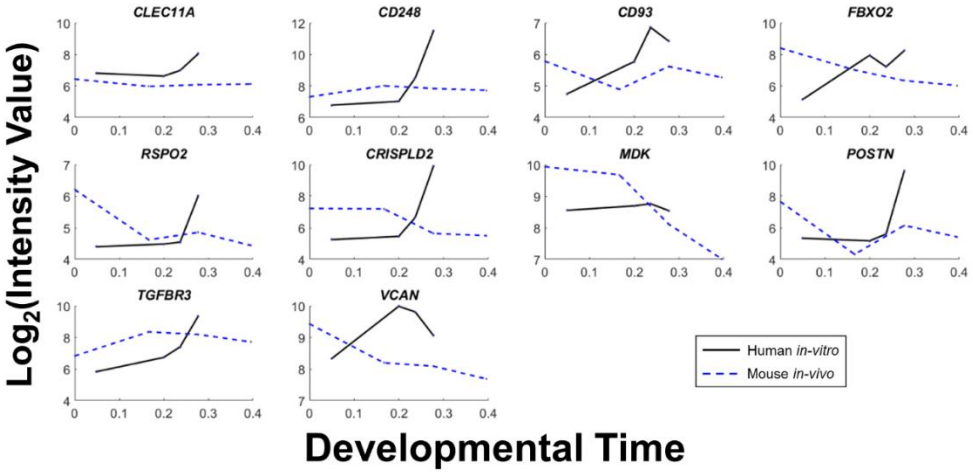

A
ADD2
LCP1
MYL4
TNNC1
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time


B
APOA4
SLC19A3
SLC1A1
SLC1A3
TDO2
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time

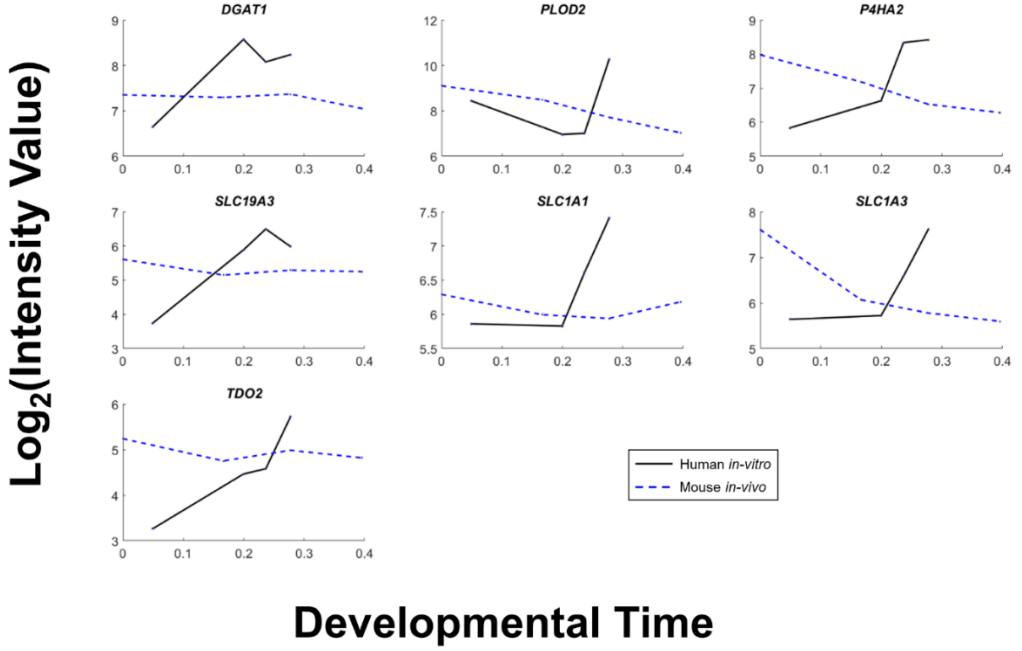

C
SLC19A3
SLC1A1
SLC1A3
TDO2
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time
D
MT1E
MT1F
MT1G
MT1H
MT1M
MT1X
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time


m
Log2(Intensity Value)
CD248
CD93
RASGRP2
S100A10
SMOC2
CDH10
CDH11
CDH13
CDH3
DSC3
FBLN1
LCP1
MMP9
MFAP4
MYL4
NPNT
NID2
PROS1
PCDH7
SPARC
SCUBE2
SLIT2
SLIT3
SYT4
THBD
TNNC1
VCAN
Human in-vitro
Mouse in-vivo
Developmental Time

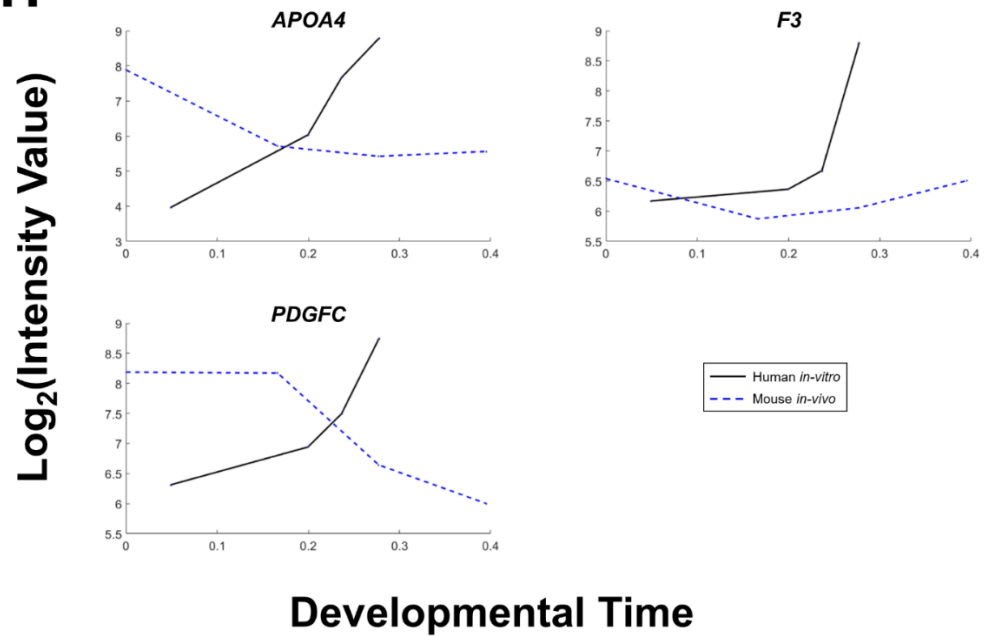
F


CLEC11A
CD248
CD93
FBXO2
RSPO2
CRISPLD2
MDK
POSTN
TGFBR3
VCAN
Log2(Intensity Value)
Human in-vitro
Mouse in-vivo
Developmental Time

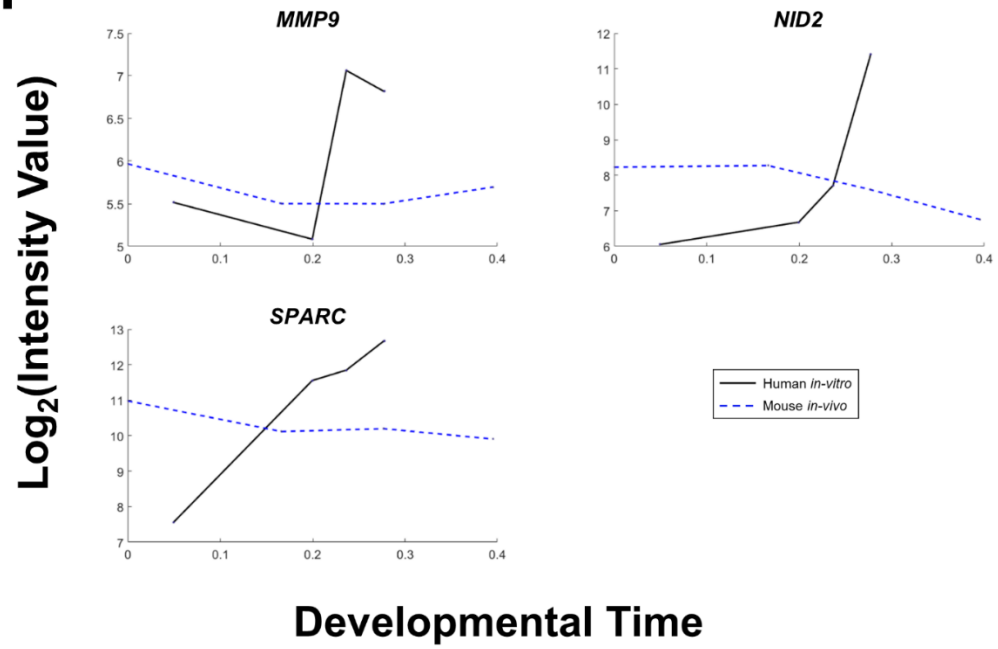
G


DGAT1
PLOD2
P4HA2
SLC19A3
SLC1A1
SLC1A3
TDO2
Log2(Intensity Value)
Human in-vitro
Mouse in-vivo
Developmental Time

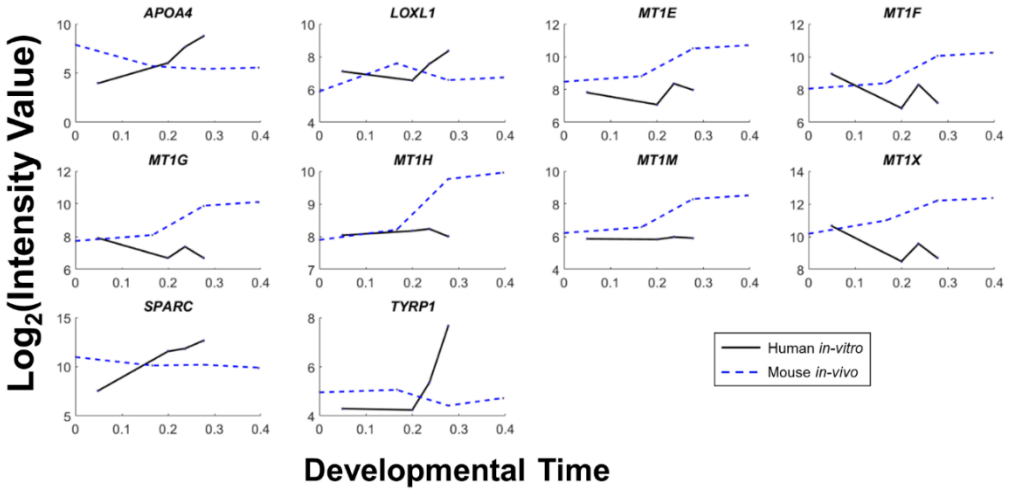
H


APOA4
F3
PDGFC
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time

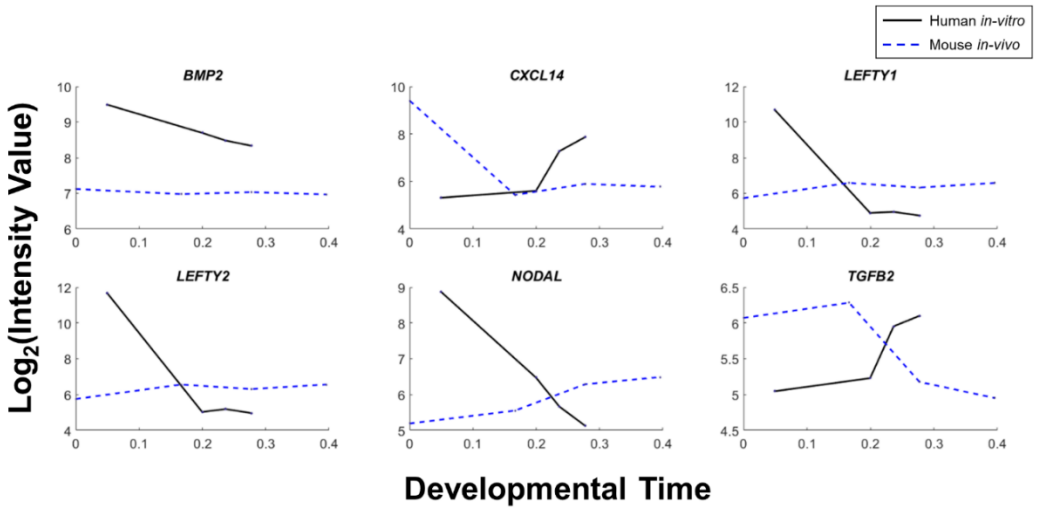
I

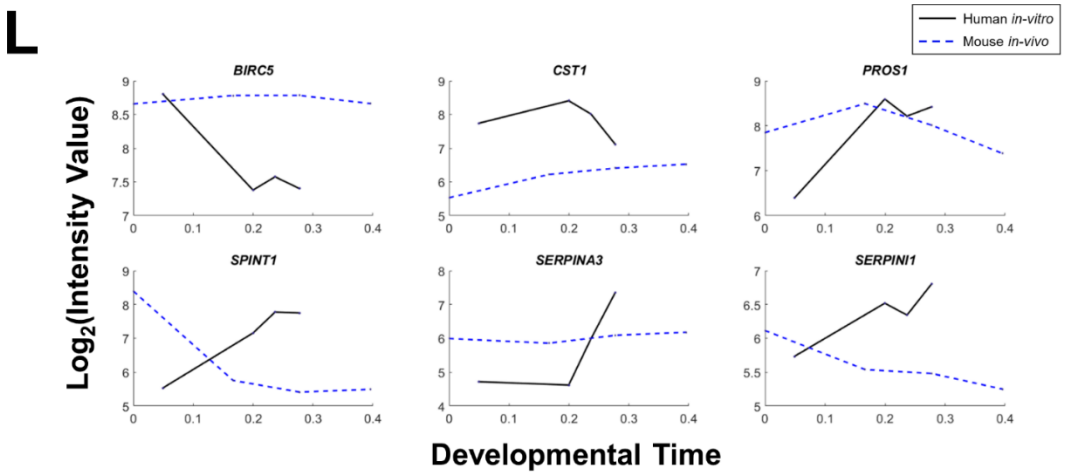

MMP9
NID2
SPARC
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time

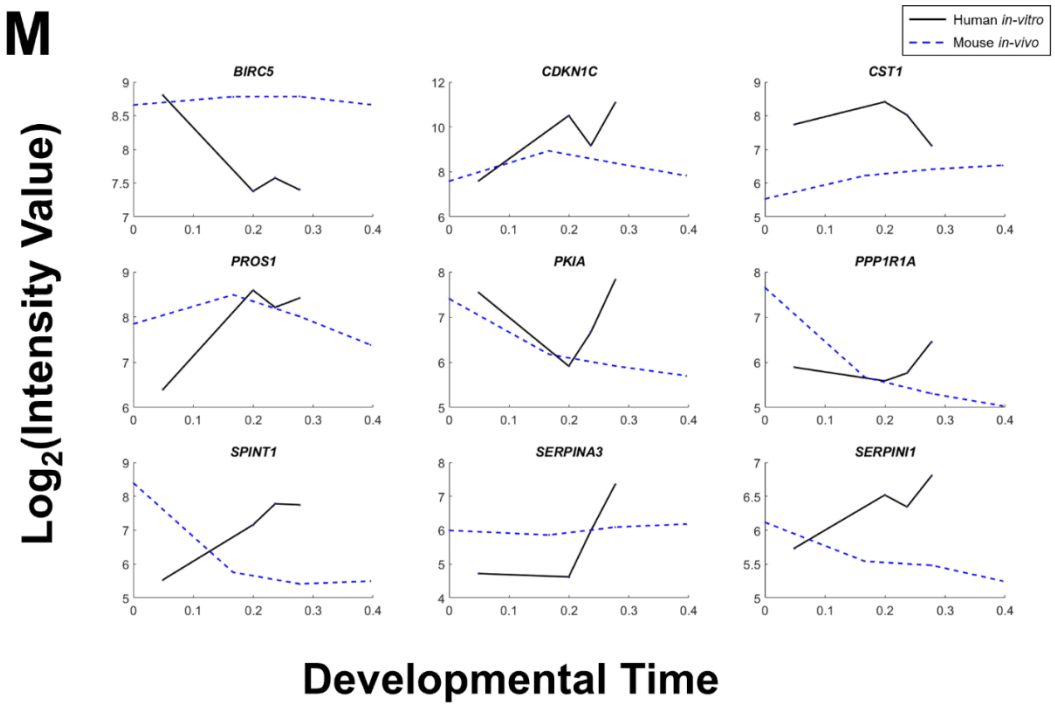
J


APOA4
LOXL1
MT1E
MT1F
MT1G
MT1H
MT1M
MT1X
SPARC
TYRP1
Log2(Intensity Value)
Human in-vitro
Mouse in-vivo
Developmental Time

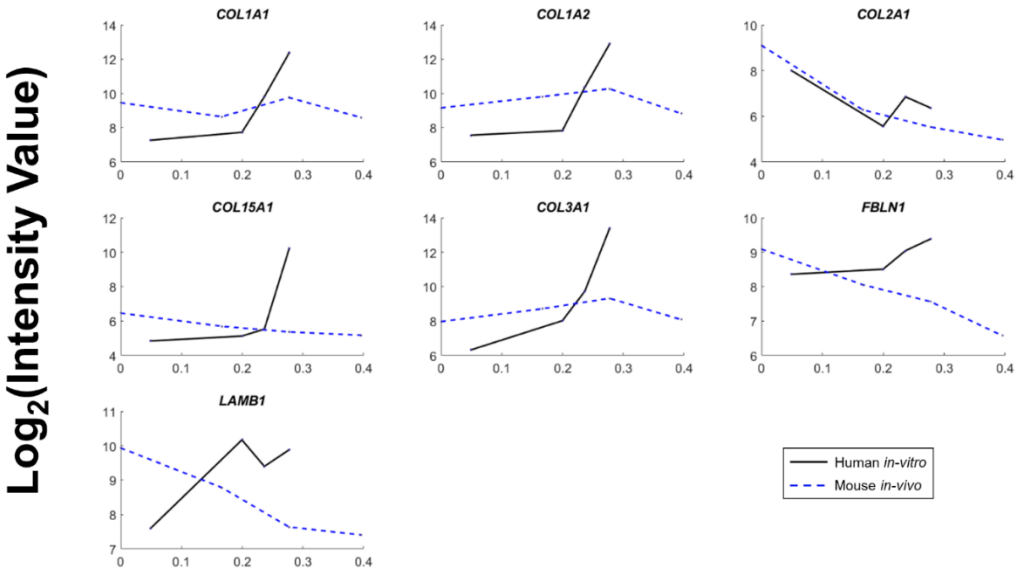
K


BMP2
CXCL14
LEFTY1
LEFTY2
NODAL
TGFB2
Log2(Intensity Value)
Human in-vitro
Mouse in-vivo
Developmental Time


L
Human in-vitro
Mouse in-vivo
BIRC5
CST1
PROS1
SPINT1
SERPINA3
SERPINI1
Log2(Intensity Value)
Developmental Time


M
Human in-vitro
Mouse in-vivo
BIRC5
CDKN1C
CST1
PROS1
PKIA
PPP1R1A
SPINT1
SERPINA3
SERPINI1
Log2(Intensity Value)
Developmental Time

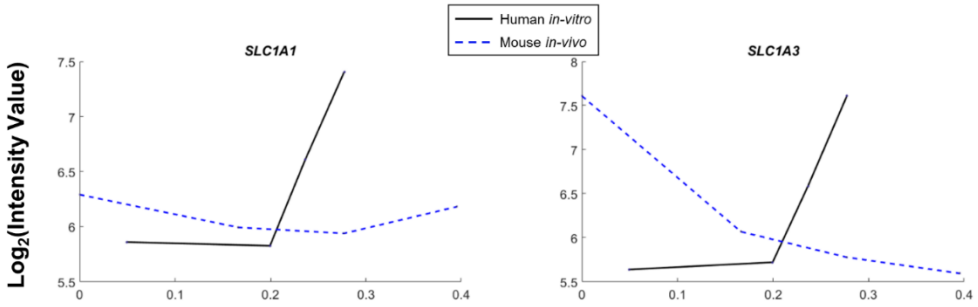
N



Log$_2$(Intensity Value)

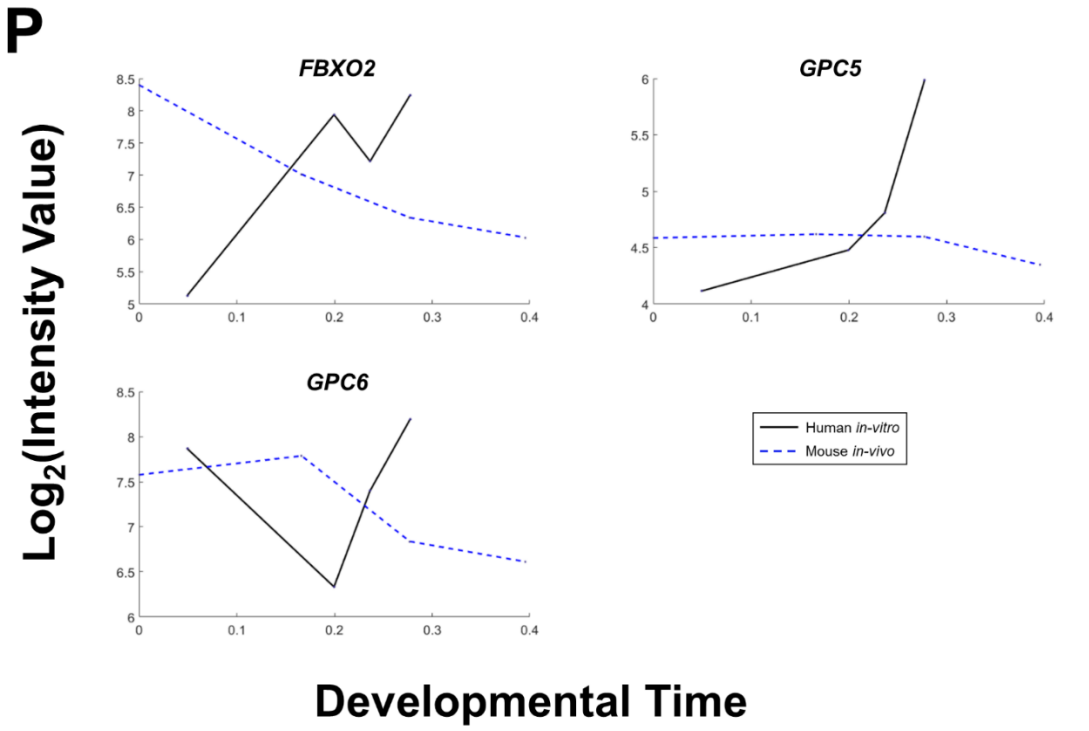
Developmental Time

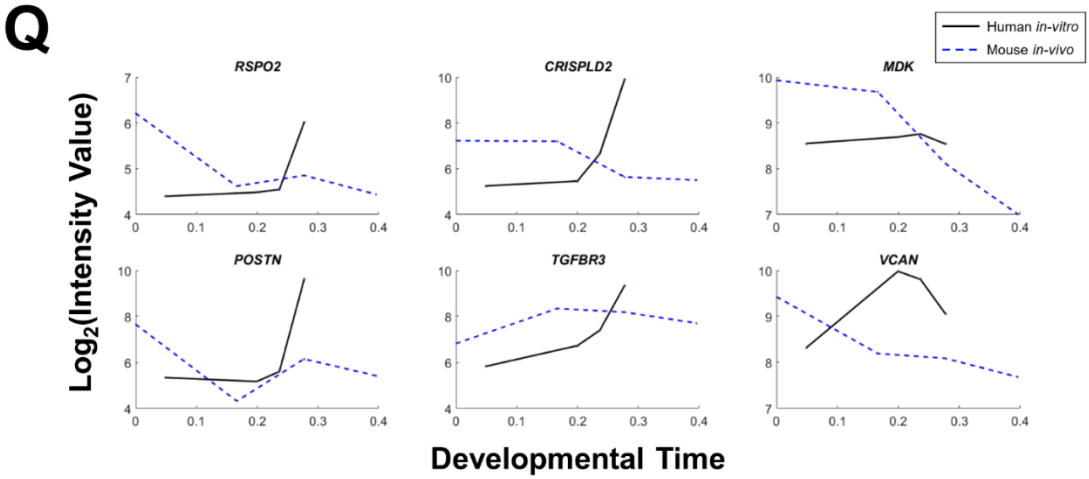
O



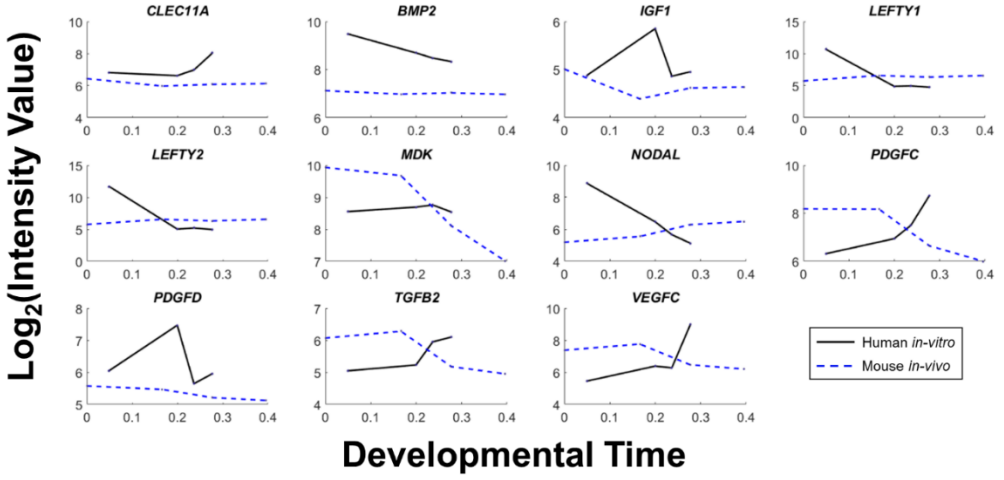
Log$_2$(Intensity Value)

Developmental Time

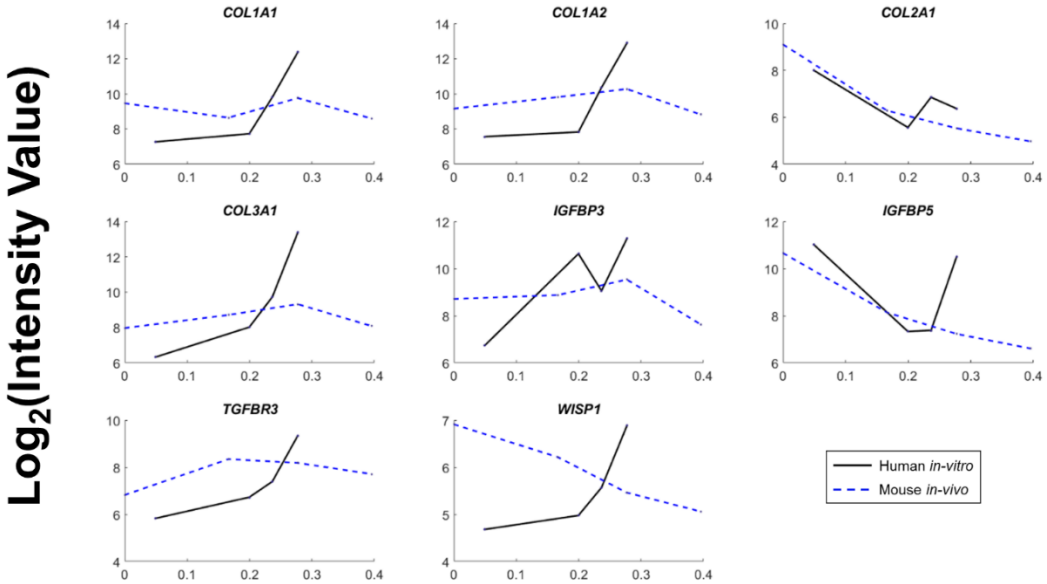

P
FBXO2
GPC5
GPC6
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time


Q
RSPO2
CRISPLD2
MDK
POSTN
TGFBR3
VCAN
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time

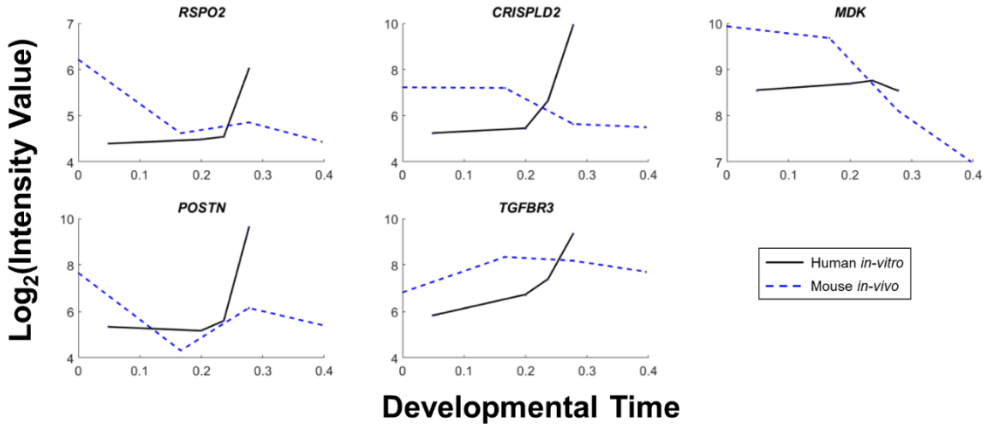
R


CLEC11A
BMP2
IGF1
LEFTY1
LEFTY2
MDK
NODAL
PDGFC
PDGFD
TGFB2
VEGFC
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time

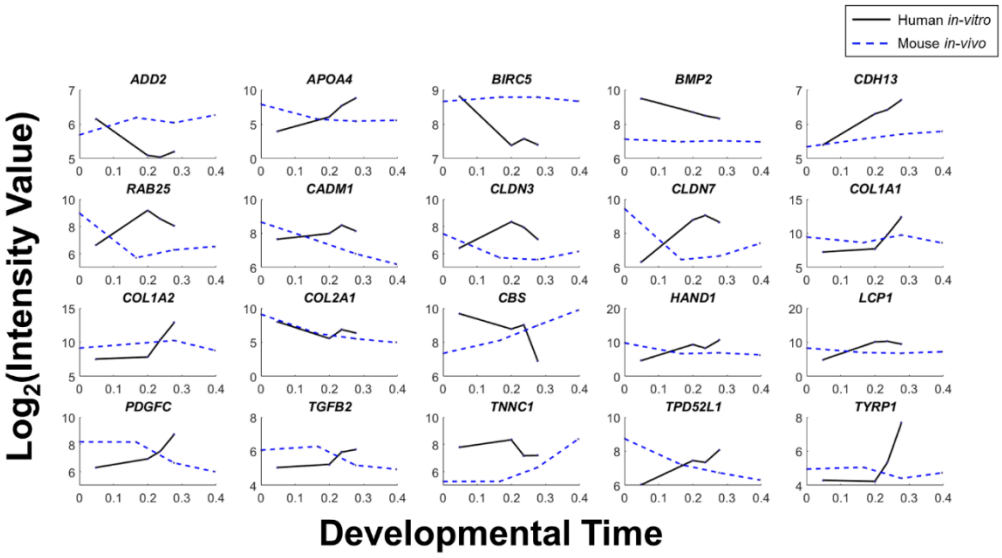
S


COL1A1
COL1A2
COL2A1
COL3A1
IGFBP3
IGFBP5
TGFBR3
WISP1
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time

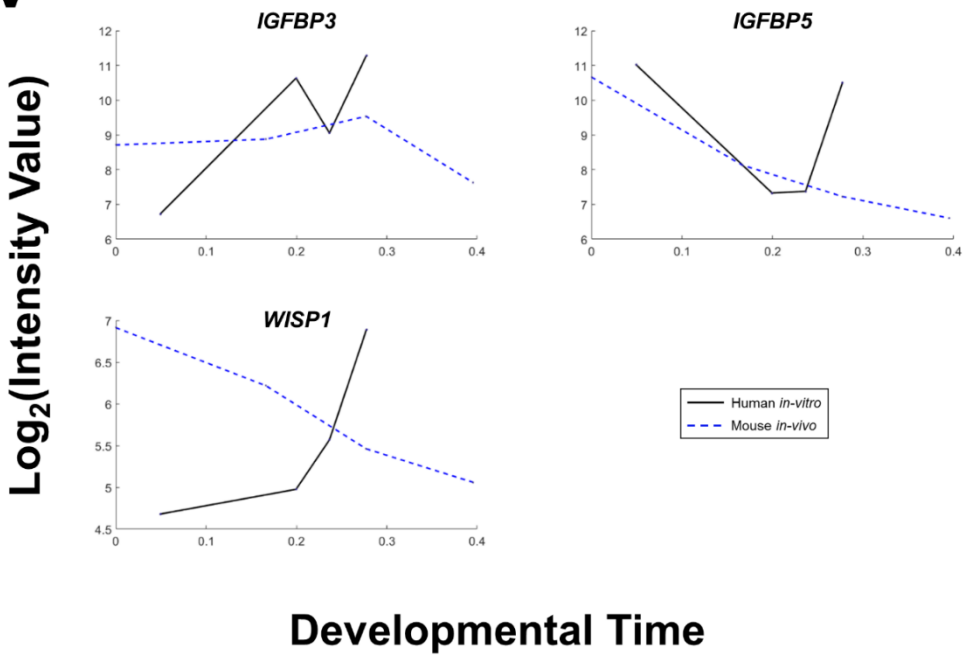
T


RSPO2
CRISPLD2
MDK
POSTN
TGFBR3
Log2(Intensity Value)
Human in-vitro
Mouse in-vivo
Developmental Time

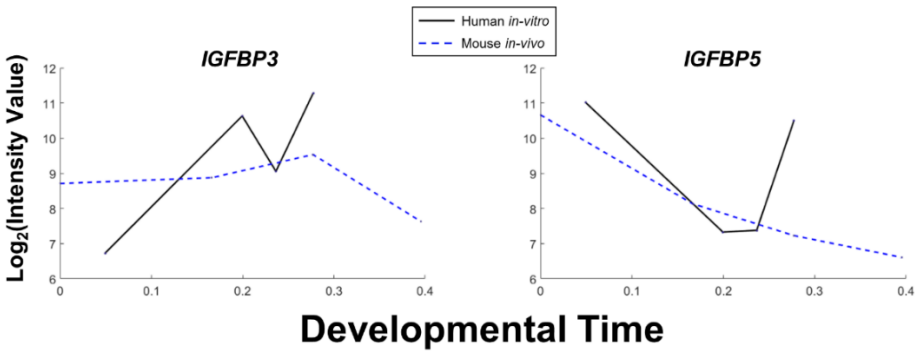
U


Human in-vitro
Mouse in-vivo
ADD2
APOA4
BIRC5
BMP2
CDH13
RAB25
CADM1
CLDN3
CLDN7
COL1A1
COL1A2
COL2A1
CBS
HAND1
LCP1
PDGFC
TGFB2
TNNC1
TPD52L1
TYRP1
Log2(Intensity Value)
Developmental Time

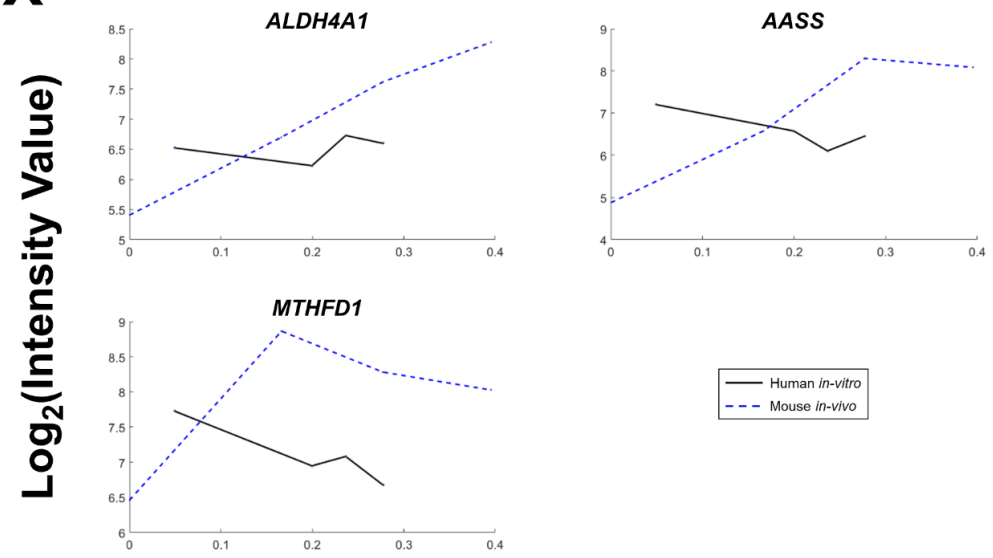
V


IGFBP3
IGFBP5
WISP1
Log2(Intensity Value)
Human in-vitro
Mouse in-vivo
Developmental Time

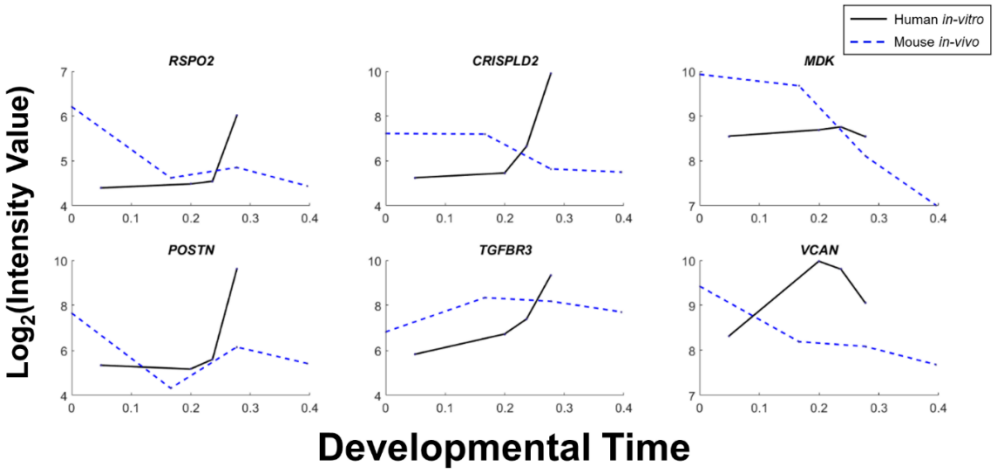
W


Human in-vitro
Mouse in-vivo
IGFBP3
IGFBP5
Log2(Intensity Value)
Developmental Time

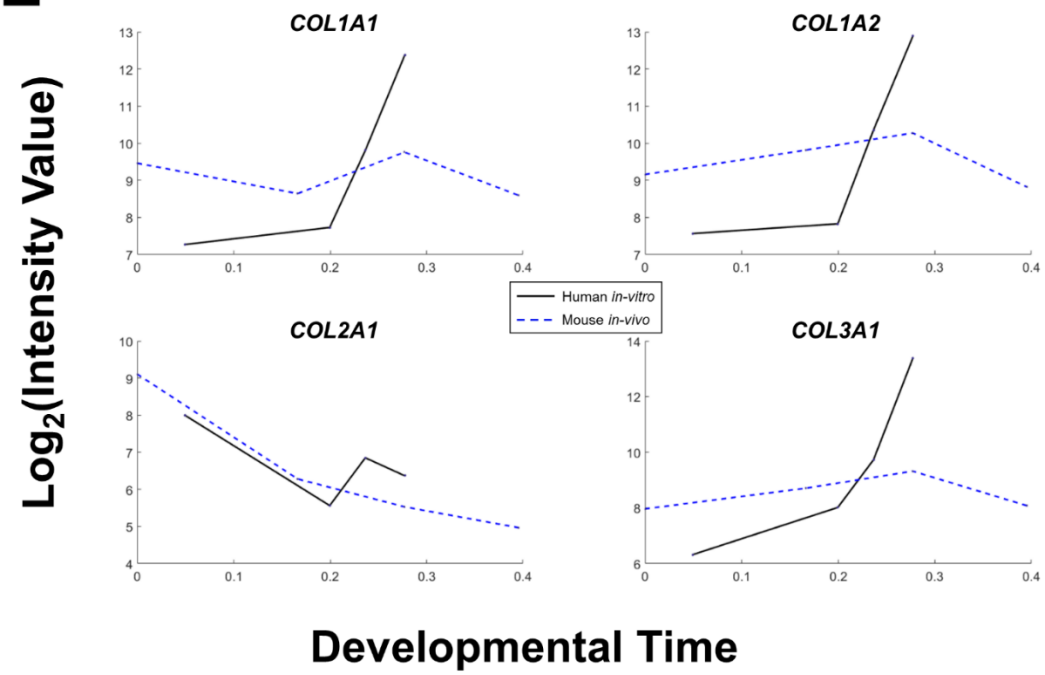
X


ALDH4A1
AASS
MTHFD1
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time

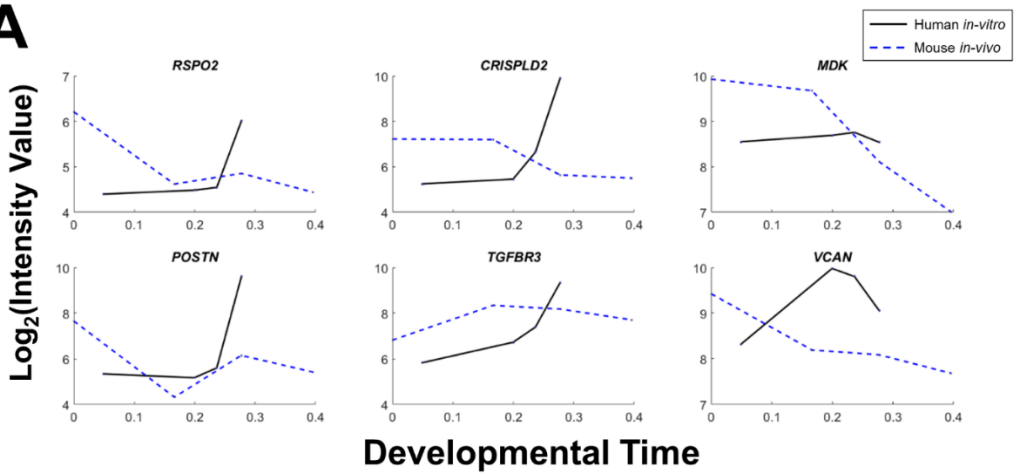
Y

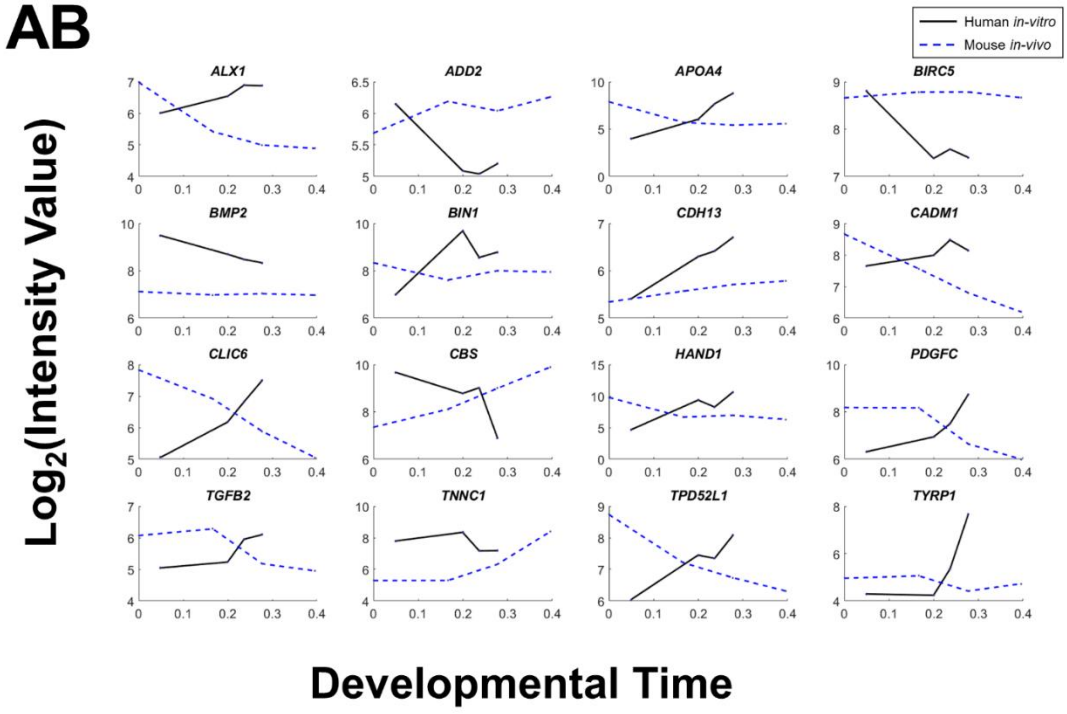

Human in-vitro
Mouse in-vivo
RSPO2
CRISPLD2
MDK
POSTN
TGFBR3
VCAN
Log2(Intensity Value)
Developmental Time

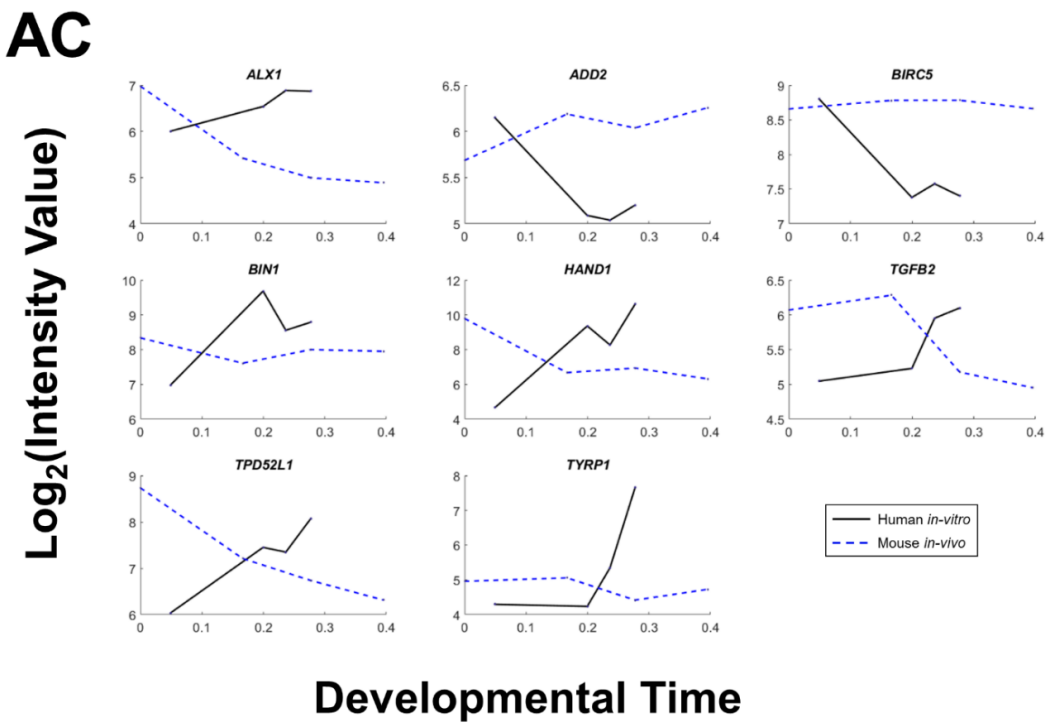
Z


COL1A1
COL1A2
COL2A1
COL3A1
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time

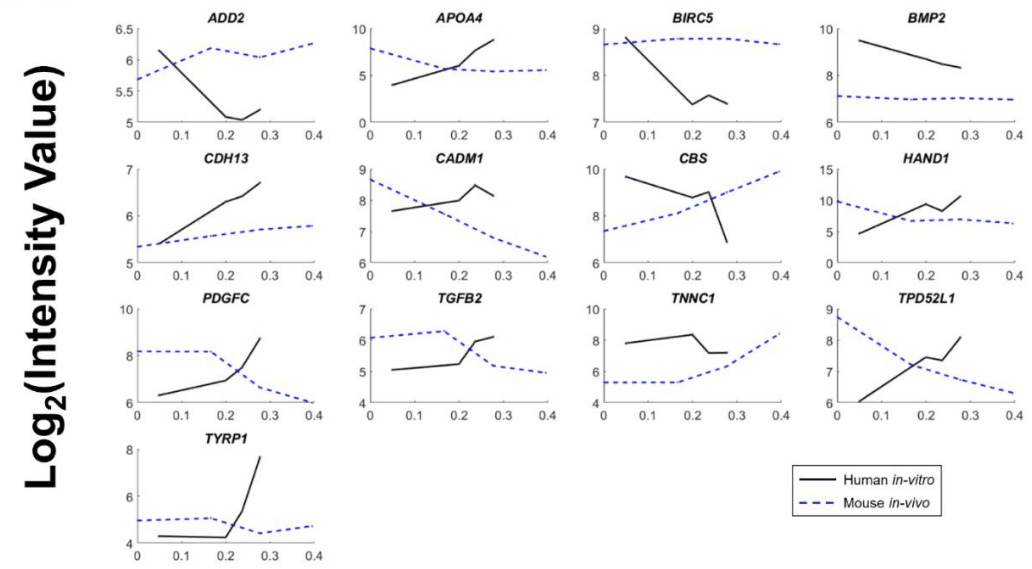
AA


RSPO2
CRISPLD2
MDK
POSTN
TGFBR3
VCAN
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time


AB
Human in-vitro
Mouse in-vivo
ALX1
ADD2
APOA4
BIRC5
BMP2
BIN1
CDH13
CADM1
CLIC6
CBS
HAND1
PDGFC
TGFB2
TNNC1
TPD52L1
TYRP1
Log2(Intensity Value)
Developmental Time


AC
ALX1
ADD2
BIRC5
BIN1
HAND1
TGFB2
TPD52L1
TYRP1
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time

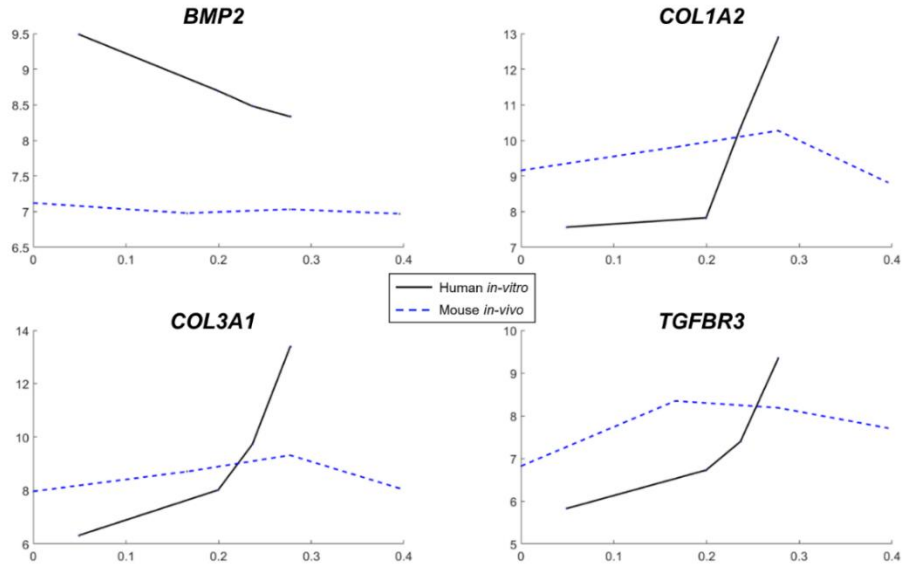
AD



Log_2(Intensity Value)

Developmental Time

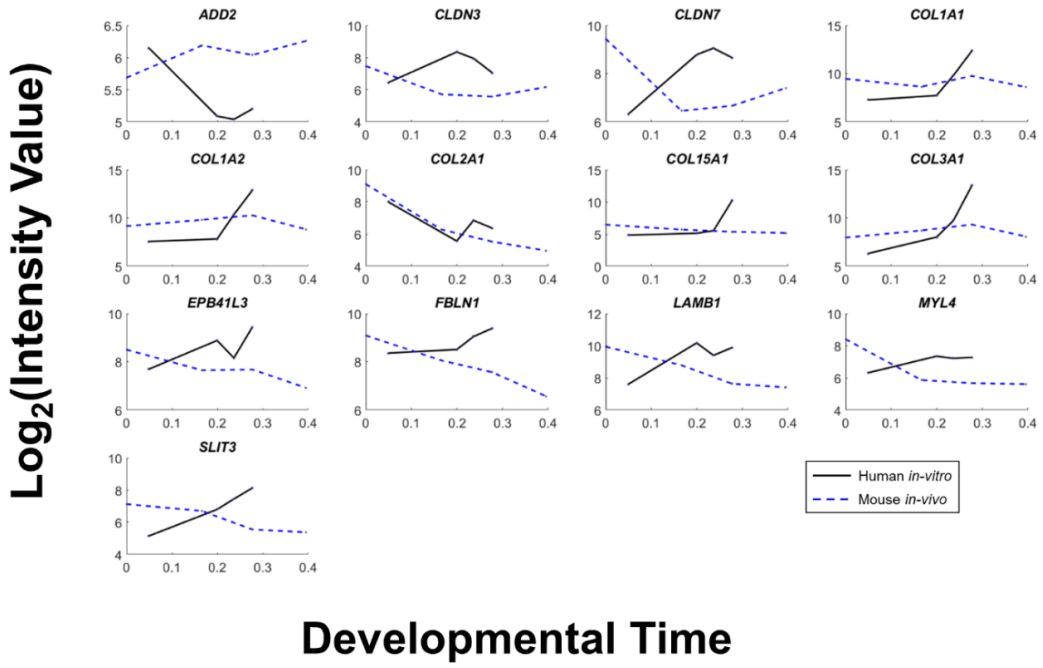
AE



Log_2(Intensity Value)

Developmental Time

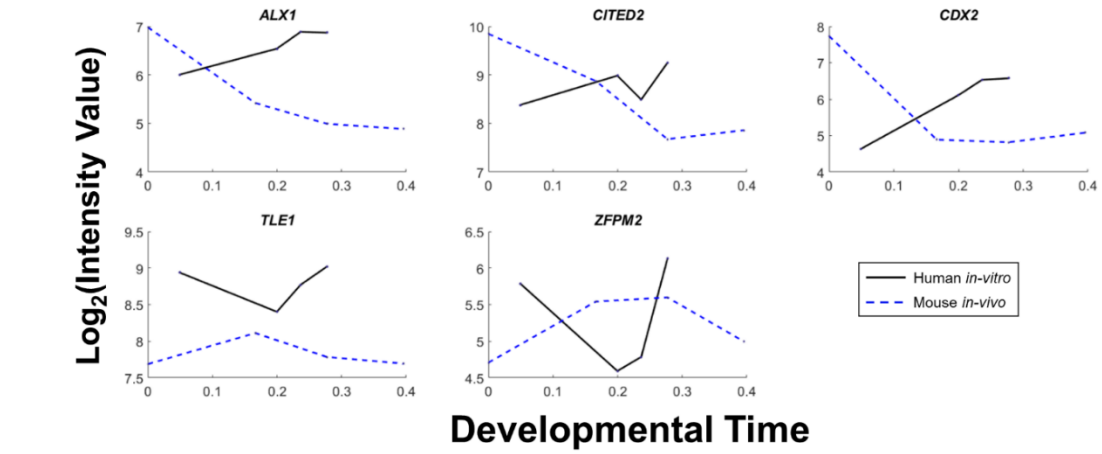
AF



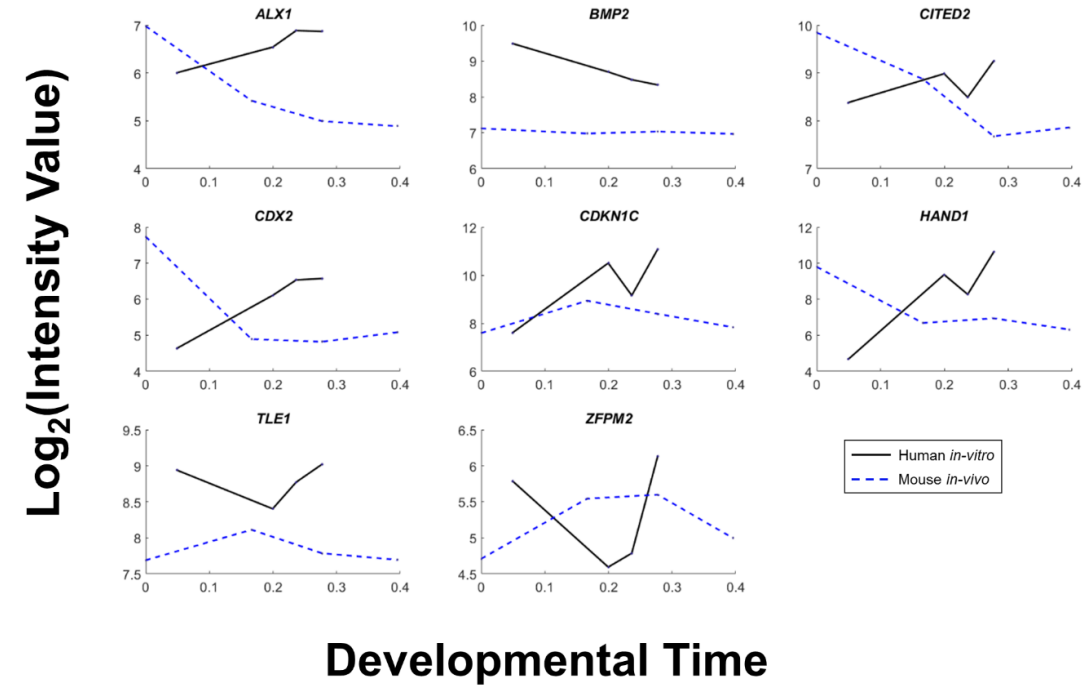
Log_2(Intensity Value)

Developmental Time

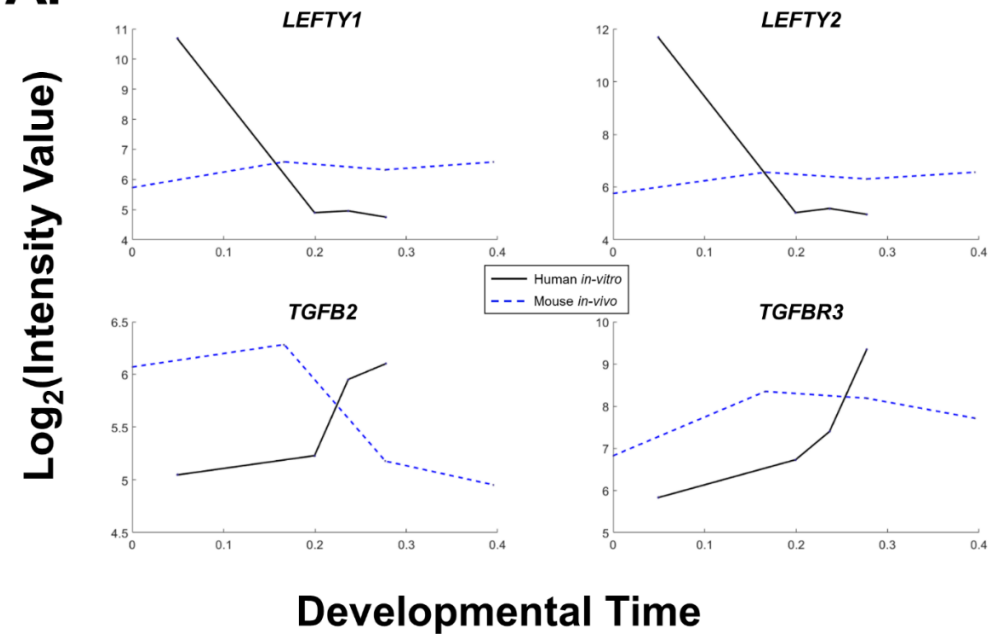
AG


ALX1
CITED2
CDX2
TLE1
ZFPM2
Log2(Intensity Value)
Human in-vitro
Mouse in-vivo
Developmental Time

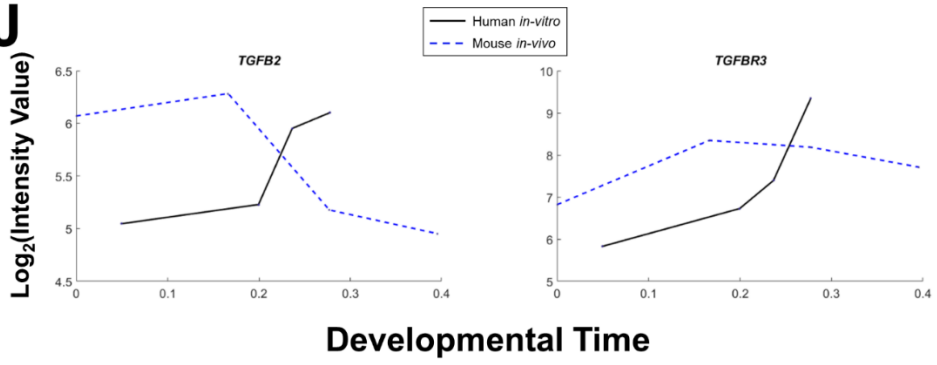
AH


ALX1
BMP2
CITED2
CDX2
CDKN1C
HAND1
TLE1
ZFPM2
Log2(Intensity Value)
Human in-vitro
Mouse in-vivo
Developmental Time

AI

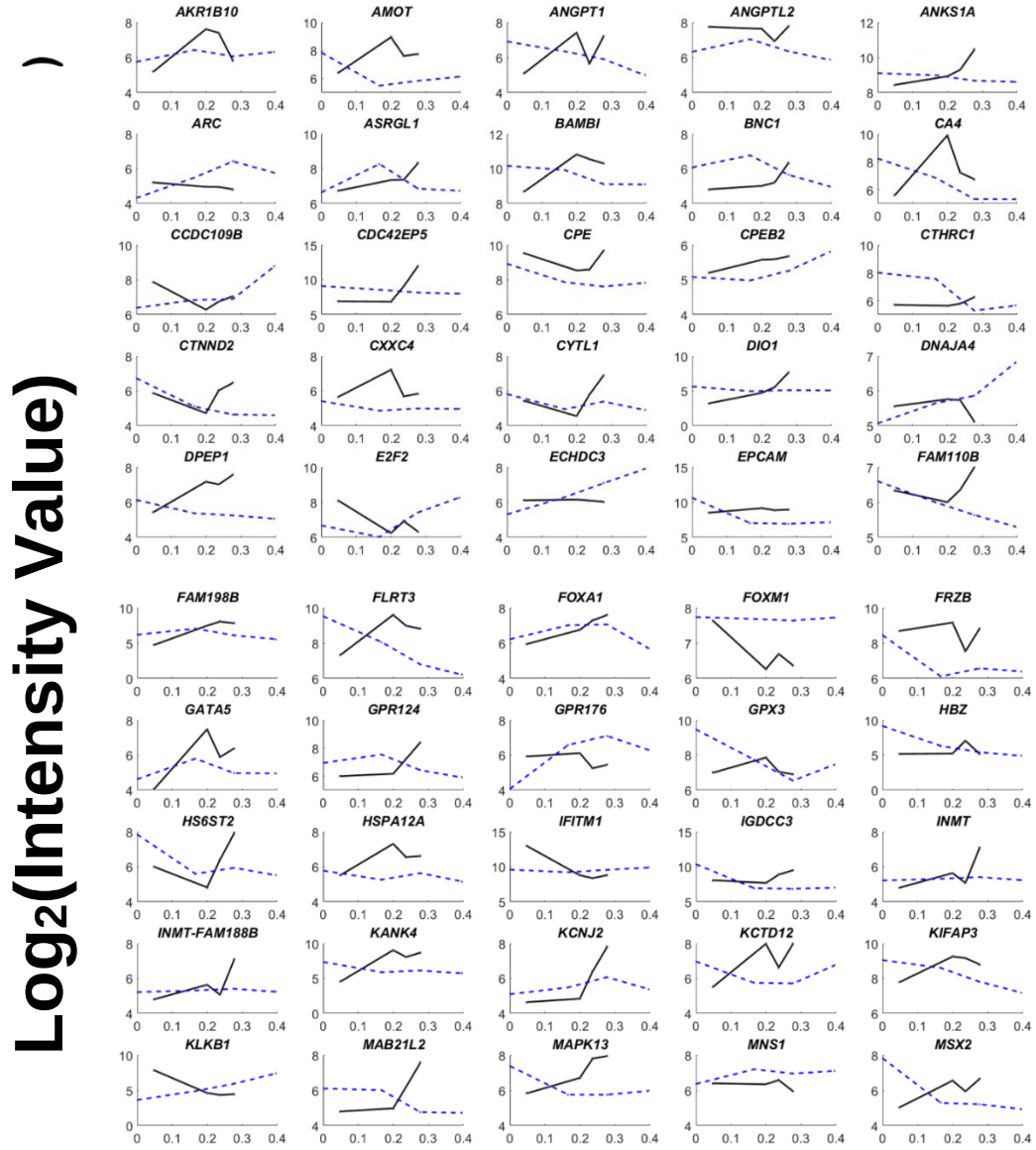

LEFTY1
LEFTY2
TGFB2
TGFBR3
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time

AJ


TGFB2
TGFBR3
Human in-vitro
Mouse in-vivo
Log2(Intensity Value)
Developmental Time

AK

)



Log2(Intensity Value)

Developmental Time

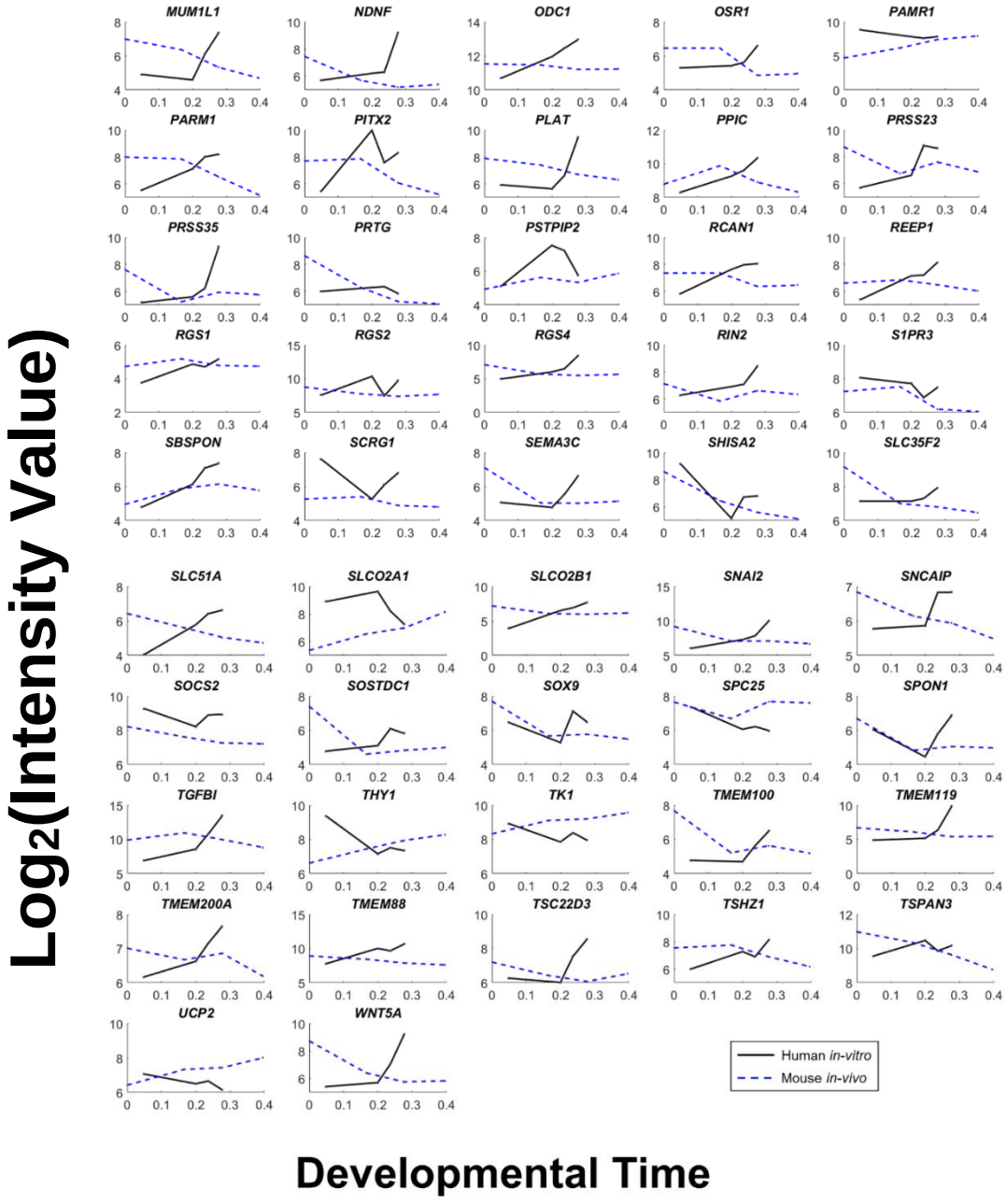


Figure 5-13A-AK. Dynamically differentially expressed genes during human *in vitro* differentiation and *in vivo* separated based on molecular function from DAVID. (Black: D6_H, D10_H, D14_H, HLC_H1) and mouse development (Blue: E9.5, E11.5, E13.5, E15.5) A) Actin filament binding B) Amine binding C) Amino acid binding D) Cadmium binding E) Calcium binding F) Carbohydrate binding G)

Carboxylic acid binding H) Cell surface binding I) Collagen binding J) Copper ion binding K) Cytokine activity L) Endopeptidase inhibitor activity M) Enzyme inhibitor activity N) Extracellular matrix protein constituent O) Glutamate binding P) Glycoprotein binding Q) Glycosaminoglycan binding R) Growth factor activity S) Growth factor binding T) Heparin binding U) Identical protein binding V) Insulin-like growth factor binding W) Insulin-like growth factor I binding X) Oxidoreductase activity Y) Pattern binding Z) Platelet-derived growth factor binding AA) Polysaccharide binding AB) Protein dimerization binding AC) Protein heterodimerization activity AD) Protein homodimerization activity AE) SMAD binding AF) Structural molecule activity AG) Transcription corepressor activity AH) Transcription represssor activity AI) Transforming growth factor beta receptor binding AJ) Type II transforming growth factor beta receptor binding AK) Unannotated

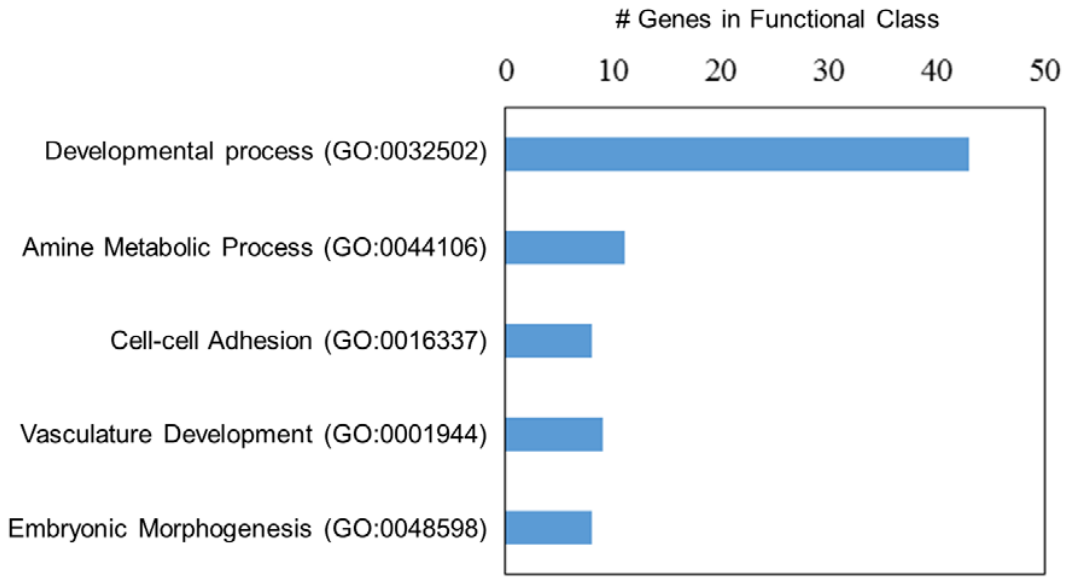


Figure 5-14. Functionally enriched gene class in the dynamically differentially expressed genes.

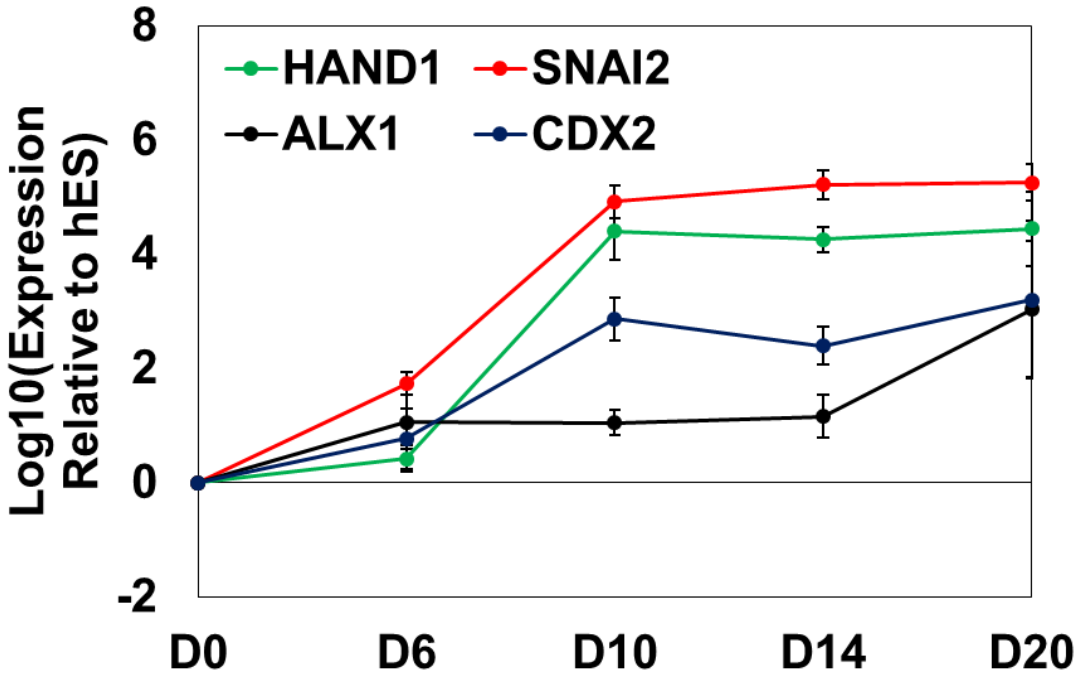


Figure 5-15. Transcript Dynamics of Notable Upregulated Transcription Factors using qRT-PCR.

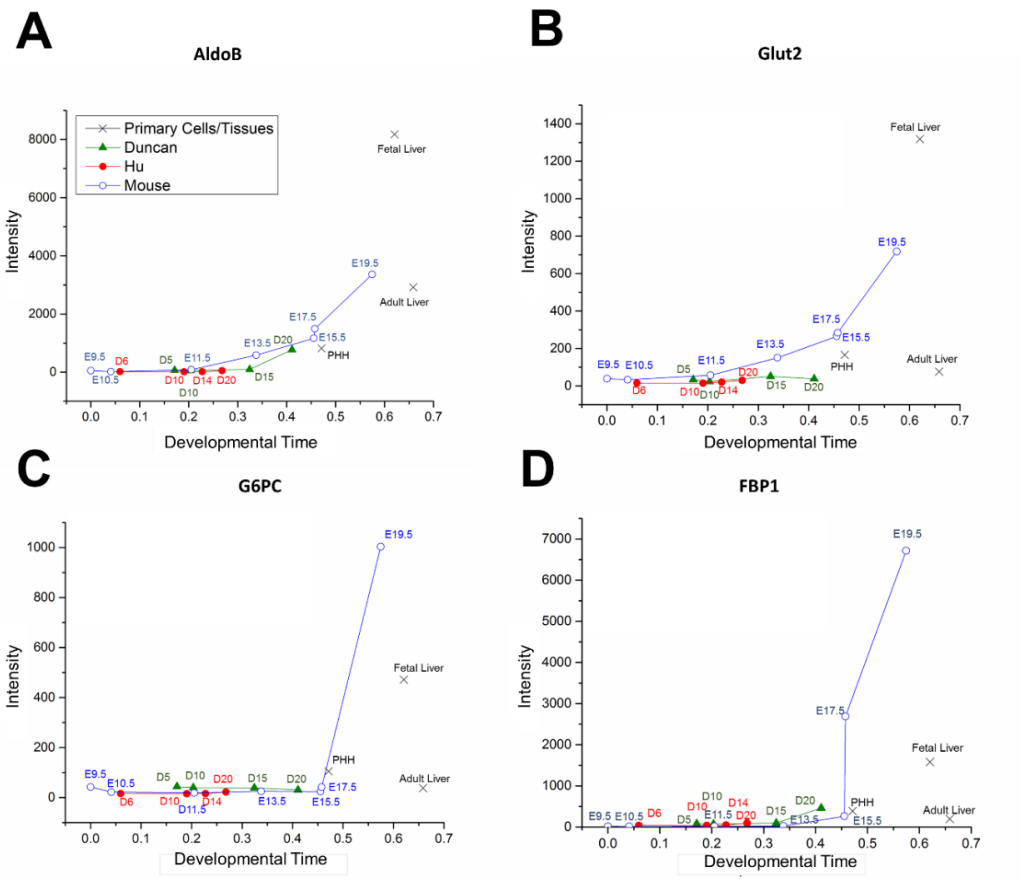


Figure 5-16. The gene expression profile for different metabolic genes. Four different genes A) aldolase B (AldoB), B) glucose transporter 2 (Glut2), C) glucose 6-phosphatase (G6PC), D) fructose-1,6-biphosphatase 1(FBP1), in glycolysis and gluconeogenesis were plotted to illustrate the difference between human *in vitro* differentiation/human liver and mouse *in vivo* development.

		Hu	Mouse
Cluster 1 $ES^{-}/HLC^{++}/PHH^{++}$	ABCA1	✓	✓
	APOA1	✓	✓
	ALDH1A1	×	×
	IL1R1	✓	✓
	CYP1A1	✓	×
	C3	×	✓
	FGA	✓	✓
	STAT4	✓	✓
	A2M	✓	✓
Cluster 2 $ES^{-}/HLC^{+}/PHH^{++}$	CYP2E1	✓	✓
	SULT1E1	✓	×
	C2	✓	×
	C6	×	×
	C9	×	×
	F5	✓	✓
	F7	✓	✓
	HPN	✓	✓
	HRG	✓	✓
Cluster 4 $ES^{+}/HLC^{-}/PHH^{--}$	CCNB1	✓	✓
	CCNB2	✓	✓
	CDK1	✓	✓
	TOP2A	✓	✓
	TTK	✓	✓
Cluster 5 ESC^{+}/HLC^{++}	COL1A1	✓	✓
	COL4A1	✓	✓
	FBN1	✓	✓
	LAMB3	×	×
	ITGA3	✓	×
	ITGA6	×	×
	EGF	×	×
	WNT5A	✓	×

Figure 5-17. Dynamic behavior of candidate genes found in (Godoy et al. 2015) were compared to our meta-analysis. Godoy et al. compared HLCs to ESCs to obtain 5 clusters of genes. We took the representative genes that Godoy presented and compared it with our dynamic data in Hu and mouse data. A check mark represents a similar behavior was found dynamically in either Hu or mouse time course data. A check mark in Cluster 1 represents a continuous increase in expression to levels similar to primary cells. A check mark in Cluster 2 represents a minimally induction of expression. A check mark in Cluster 4 represents a continuous decrease in expression to levels similar to primary cells. A check mark in Cluster 5 represents an increase in expression level from an existing high expression level.

Table 5-1. List of microarray transcriptome data of human HLC differentiation complied for this study.

Data	Platform	Samples	Brief Description	Abbreviation of Samples
GSE14897 (Duncan)	Affymetrix	H9: D0, D20 iPS: Fibroblasts, D20	Hepatocyte-like cells from iPS and H9 differentiation on D20	HLC_D, iHLC_D
GSE25417 (Duncan)	Affymetrix	H9:D5, D10, D15, HNF4 knockdown D20	HNF4α knockdown in H9 differentiation	D5_D, D10_D, D15_D
GSE25744	Illumina	H1: D0, HLC iPS: D0, HLC fibroblasts, fetal liver	H1 and iPS cells differentiated using 2 protocols	HLC_A1, HLC_A2, HLC_A3, HLC_A4, HLC_A5, Fetal Liver
	Illumina	H9: D0, D6, D9, D10, D12, D14, D20, D26	Time-course of H9 cells differentiation of HLCs	D6_H, D9_HE, D10_H,D12_HE,D14_H, HLC_H1,HLC_H2
Raju et al, to be deposited	Illumina	HSF6: D0, D20, D32 spheroids, Adult liver	Hepatocyte-like cells from HSF6 cells differentiated to D20 and cultured as spheroids as well as adult liver tissue	HLC_S,HLC_SS,Adult Liver
Raju et al, To be deposited	Affymetrix	HES: D0, D20	Hepatocyte-like cells from HES3 differentiation	HLC_H3

GSE42643	Agilent	Fibroblasts, HLCs, PHH	Hepatocyte-like cells reprogrammed from human fibroblasts and primary hepatocytes	iHEP,PHH
GSE46631	Agilent	Liver bud from iPS	Liver bud: Hepatocyte-like cells from iPS with HUVECs and MSCs	LB

Table 5-2. List of mouse microarray transcriptome data compiled in this study.

Data	Platform	Samples	Brief Description	Abbreviation of Samples
GSE46631	Agilent	E9.5 - E19.5	Mouse liver development	E9.5, E10.5, E11.5, E13.5, E15.5, E17.5, E19.5
GSE13149	Affymetrix	E11.5 - Adult	Mouse liver development	E11.5, E12.5, E13.5, E14.5, E15.6, E16.5, E17.5, E18.5
www.alexaplatform.org	Illumin HiSeq	E8.5, E14.5, Adult Liver	Mouse liver development	E8.5, E14.5, PHH_ms
GSE23635	Illumina	MEFs, iHEP	MEFs reprogrammed to hepatocytes	iHEP1
GSE29725	Illumina	MEFs, iHEP	MEFs reprogrammed to hepatocytes	iHEP2

GSE48486	Affymetrix	MEFS, iHEPSCs	MEFS reprogrammed to hepatic progenitor cells	iHEP3
To be deposited	Illumina	D6, D10, D20, D28	miPSCs differentiation to HLCs	D6, D10, D20, D28

Table 5-3. Transcriptional factors overrepresented in differentially expressed genes based on their predicted binding sites.

Overrepresented TFs	Number of Genes Dynamically Differentially Expressed
PPARG	95
AP4	79
RP58	69
ER	65
TATA	65
GR	64
AHRARNT	63
GATA	56
CHX10	51
IK1	50
LYF1	49
ZIC2	31
GATA6	26
CREL	23

Chapter 6. Metabolism Shift in Hepatocyte-like Cells

6.1 Introduction

In the previous chapter, we observed significant changes in numerous genes over the course of the differentiation. Some of the most obvious changes that occurred were involved in the metabolic pathways. In some incidents, certain genes that are involved in mature hepatic functions that are involved with glucose metabolism, however, was found to never be turned on when compared to their primary hepatocytes counterpart. In this chapter, we dived a bit further to illustrate if our hESCs are transitioning towards a metabolic phenotype that mimics primary hepatocytes. The understanding of metabolism in mammalian cells has contributed to maximizing the capability of using cells as the major workhorse in the biotechnology industry and medicine field. Our increasing understanding on the effects of metabolism on cellular behavior has resulted in massive efforts to designing robust cell culture practices for reproducible production of different biologics. Over the past decades, as our knowledge in stem cell and developmental biology grows, stem cells have emerged as the next cellular candidate for revolutionizing the field of regenerative medicine and tissue engineering.

Stem cells possess the ability to self-renew and differentiate into highly specialized cell types found in the body. Using different growth factors and cytokines to mimic embryonic signaling cues, stem cells can be guided to differentiate into hepatocytes through different developmental stages in a manner similar to embryonic development. Thus, a better understanding of the different mechanisms controlling expansion and differentiation of stem cells is of great importance for the systematic production of cells for clinical and industrial applications. Various analyses have shown the resemblance of stem cell-derived hepatocytes (SCDHs) to primary hepatocytes on different levels, particularly at the global gene expression and proteomics. Although the cellular metabolism has been looked at extensively, the role of cellular metabolism has on hepatocyte differentiation is not as well established.

Primary hepatocytes have been shown to rely primarily on mitochondrial respiration for bioenergetic demands. The generation of stem cell-derived functional and

metabolically competent hepatocytes would be a key step towards the utilization of SCDHs for regenerative medicine in liver applications. It is well documented that each type of cell has a unique metabolic profile to dictate the function or fate of that cell type. During differentiation, stem cells transition through several stages undergoing genetic and biochemical changes to fit the energetic demand of the current cell type. Stem cells and other fast growing cell types (e.g. CHO, cancerous cell types), have long been observed to rely on glycolysis for energy. In agreement with these preferences for energy demand, hypoxia has been shown to be beneficial in maintaining hESCs in its pluripotent state (163, 164). In more differentiated cell types, such as primary hepatocytes, mature mitochondria are typically observed with the activation of the mitochondria process of oxidative phosphorylation(OxPhos) for energy supply (165). Studies have shown that primary hepatocytes exist in a low glycolytic state with a metabolic signature representative of oxidative phosphorylation with moderate levels of oxygen consumption(166, 167). The preference between glycolysis and oxidative phosphorylation have often been used as a marker of differentiation status. The gene expression profile of pluripotent stem cells reveal slightly different metabolic signature compared to their differentiated counterpart in terms of glucose-dependent pathways (168).

While there is plenty of studies on the metabolic behavior of pluripotent stem cells during differentiation, the behavior of primary hepatocytes and pluripotent stem cells have been studied much more extensively. Pluripotent stem cells have been shown to exist in a high state of glycolysis with high glucose consumption and lactate production. However, freshly isolated primary hepatocytes in culture are much more metabolically active capable of drug metabolism and gluconeogenesis. Studies have shown that primary hepatocytes exist in a low glycolytic state with a metabolic signature representative of oxidative phosphorylation with moderate levels of oxygen consumption.

Few studies have assessed the bioenergetic changes in pluripotent energy metabolism during differentiation towards the hepatic lineage. Due to the dynamic nature of stem cells during differentiation, it is important for us to understand the bioenergetic changes that is occurring. During the very early stages of embryo development, cells experience a metabolic shift from oxidative phosphorylation to glycolysis, and reverts back

to oxidative phosphorylation after commitment to tissue lineage (169). For example, the restriction or abundance of certain amino acids can alter the metabolic behavior and lead to certain byproducts changing the cell's fate and function.

In accordance to previous studies, our results reveal that hESCs have mitochondria with lower activity consistent with cells in low oxidative phosphorylation and high glycolytic activity. In this study, we looked at the changes occurring during differentiation towards the hepatic lineage and the changes in bioenergetic demands. In addition, the effect of glucose availability on stem cell proliferation and differentiation was examined closely in this work. A comprehensive analysis using transcriptome and metabolic flux model was used to shed some light on the mechanisms that hESCs use to maintain high levels of glycolysis and the changes that occurring during differentiation.

6.2 Results

6.2.1 Hepatocyte Differentiation from Human Embryonic Stem Cells on Microcarriers

A 20-day differentiation protocol previously published before was used to differentiate H9 human embryonic stem cells towards hepatocyte-like cells (170). We used the same protocol to differentiate feeder-free H9 embryonic stem cells towards the hepatic lineage. As the differentiation proceeded, we assessed the level of stage-specific transcripts during the differentiation using qRT-PCR. Endodermal markers, *goosecoid* (*GSC*) and CXC motif chemokine receptor 4 (*CXCR4*), were both expressed in the second stage of hepatic differentiation as reported previously. We also observed the gradual increase in expression of different hepatic markers; alpha-fetoprotein (*AFP*), albumin (*ALB*), Alpha1-antitrypsin (*AAT*), and hepatic nuclear factor 4 (*HNF4α*), reaching the highest level by the end of the differentiation.

In line with the transcript levels, we also evaluated the expression of endodermal and hepatic genes at the protein level using immunostaining. FOXA2, a key transcription factor in the establishment of endodermal cells, were prominent in the second stage consistent with the transcript levels observed. We also observed the presence of HNF4α, a nuclear transcription factor expressed during liver development which regulates the

expression of several liver-specific genes, by the second stage of differentiation. At the end of the differentiation, we observed a majority of cells expressing albumin. These results suggest that we can successfully differentiate embryonic stem cells to hepatocytes after the exposure to different cytokines and growth factors over twenty days to similar reports.

To increase the volumetric consumption rate of nutrients and amino acids, microcarriers were utilized in this study to obtain higher cell concentration for measurable differences in metabolite concentrations between consecutive samples (Figure 6-1A). Numerous reports have shown the difficulty in attaching embryonic stem cells onto different configurations of microcarriers so we decided to use differentiating endodermal cells. From our previously published study, we observed that endodermal cells are capable of proliferating to a large extent with the capability to still differentiate towards the hepatic lineage. We screened different coating material on different microcarriers and found Matrigel™-coated Cytodex 3 microcarriers to have the highest attachment efficiency (Figure 6-2). Thus, we assessed the extent of differentiation for both method of differentiation on Cytodex 3 beads and tissue culture plates (TC). To determine the progression through differentiation in both methods, we investigated the presence of different markers using immunohistochemistry and qRT-PCR.

Using confocal microscopy for Cytodex 3, we observed a majority of cells expressed the endodermal marker, FOXA2, on Cytodex 3 and tissue culture plate (TCP) differentiation after the endodermal induction media. Similarly, the presence of the liver transcription factor, HNF4α, was expressed in the second stage of differentiation illustrating a similar progression of differentiation for both Cytodex 3 and TCP. By the end of the differentiation protocol, we observed that a majority of the cells were positive for mature hepatic marker, albumin (ALB), on both Cytodex 3 and TCP differentiation (Figure 6-1C and 6-3A).

As visualized by the presence of certain stage markers, qRT-PCR using primers specific to different hepatic markers were used to measure the transcript level at the end of the 20-day protocol. In the present study, hepatic markers (ALB, AAT, ASGPR1, and CYP3A4) were observed to be significantly expressed by the end of the differentiation

process (Figure 6-1B). The transcript level revealed that the HLCs differentiated on Cytodex 3 and TC showed remarkably similar levels of different hepatic markers.

Functional maturity of the hepatocyte-like cells obtained from both differentiation methods were also assessed by measuring albumin and urea synthesis (Figure 6-3B and 6-3C). Both analysis showed HLCs capable of secreting albumin and urea that is typically observed in primary hepatocytes in the liver. The levels of secretion for both albumin and urea were comparable to previously reported rates as well (78, 170).

6.2.2 Metabolic Switch During Hepatic Differentiation

Representative data for different metabolite (glucose, oxygen, and lactate) were monitored in the medium at different days of the differentiation for calculating the specific rates (Figure 6-4). During HLC differentiation, glucose consumption on a per cell basis (i.e. specific glucose consumption rate) decreased progressively over the course of the differentiation. Accompanied by the decrease in glucose consumption, we also observed a general decrease in lactate production. The maximum theoretical yield of lactate production from glucose consumption in glycolysis is 2.0. The molar stoichiometric ratio of lactate production to glucose consumption ($\Delta L/\Delta G$) decreased from 1.30 ± 0.35 (mol/mol) during the first stage of differentiation to 0.50 ± 0.18 (mol/mol) by the second stage of differentiation. By the end of the differentiation, the derived hepatocyte-like cells exhibited a $\Delta L/\Delta G$ of 0.26 ± 0.06 indicative of cells in a low glycolytic state (Figure 6-5A). The consumption rate for most the amino acids decreased over the course of differentiation. Levels of alanine, glutamate and proline increased in the media over the course of the differentiation indicating secretion of these non-essential amino acids.

Oxygen consumption increased in parallel to the decrease of glycolytic activity over the course of differentiation, from $1.50*10^{-10}$ mmol O_2/cell/h initially to $2.29*10^{-10}$ mmol O_2/cell/h by the end of the differentiation (Figure 6-5B). The stoichiometric ratio of oxygen and glucose consumption increased from 1.21 ± 0.20 (mol/mol) to 3.46 ± 0.41 (mol/mol) by the end of the differentiation. The data suggests that cells adopt an increased OxPhos metabolism as they mature further into hepatocyte-like cells.

6.2.3 Metabolic State Change Reflected at Transcript Level

To further investigate the metabolic pathways that may be responsible for the metabolic switch, we looked at the transcript levels of key glycolysis genes involved in the regulation of glucose. Although we did not observe drastic changes in many key glycolytic enzymes from microarray data (Figure 6-6), we did observe the biggest changes in the glucose transporter (GLUT), hexokinase (HK), and pyruvate kinase (PKM). Using the microarray data, we further narrowed down glycolytic enzymes to the most highly expressed isoforms in our cells and carried out qRT-PCR to confirm our findings from the microarray data. We observed a substantial decrease in expression of GLUT3 for HLCs compared to hESCs and over 2-fold decrease in expression for GLUT1 by the second stage. In addition, we also saw a similar trend in gradual downregulation of PKM1/2 and HK1 by almost 2-fold after the first stage. The collective set of genes showing a gradual decrease of expression over the course of differentiation suggest that there may be master regulator driving their behavior. Many reports have suggested that the AKT pathway can affect cellular growth and metabolism through the regulation of the different glycolytic enzymes that were mentioned previously (75). However, no such changes were seen at the transcript level for the AKT gene and Western blot also revealed no changes on the protein level when performed on Total AKT and phosphorylated AKT (Figure 6-7). Surprisingly, we did see a dramatic increase in expression for mitochondria pyruvate carrier (MPC1/2) from the microarray data (Figure 6-8). Using qRT-PCR, we confirmed our observation by observing an increase of 4-fold by the second stage in expression compared to the starting stem cell state.

6.2.4 Mitochondrial Morphology During Differentiation along the Hepatic Lineage

Changes in mitochondrial morphology and activity have been reported to indicate a preference towards cells with an oxidative phosphorylation metabolism. The observed transition between glycolysis and oxidative phosphorylation during the hepatocyte differentiation was further characterized by looking at the morphology and activity of the mitochondria over the course of differentiation. Using MitoTracker Green FM which stains the mitochondria regardless of membrane potential, we used confocal microscopy to show

that the mitochondrial population was much less organized and dimmer in intensity (Figure 6-9A). However, when we looked at a differentiated population under the same imaging settings, we saw a brighter network of elongated mitochondria typically observed in mature mitochondria. The morphology in mitochondria has been frequently reported to change in accordance to the metabolic needs of cells [insert citation]. We assessed mitochondrial morphology by using a quantitative analysis measuring mitochondrial form factor and aspect ratio during over the course of differentiation. We showed that mitochondrial elongation (aspect ratio) and network formation (form factor) showed an increase by stage 4 of differentiation consistent with the observation in increase in OxPhos (Figure 6-9B).

6.2.5 Mitochondrial Biogenesis and Activity During Hepatocyte Differentiation

In addition to assessing the mass of mitochondria in our cells over the course of differentiation, we also considered mitochondria activity by using MitoTracker Red CM-h2xros which fluoresces upon activation of membrane potential. We observed a much lower membrane potential as illustrated by Figure 6-10A in the beginning of differentiation. Cells have much more active mitochondria by the end of the differentiation starting in the second stage of differentiation corresponding to the increase in oxygen demand found previously. Due to the observation of increase in mitochondria mass and activity, we next analyzed the mitochondria DNA in cells over the course of differentiation. These results were further verified by showing an increase in mitochondria DNA levels over the course of differentiation (Figure 6-10B). We revealed that cells doubled their mtDNA levels from $0.015 \pm 0.002 \times 10^{-12}$g/cell in the beginning to $0.034 \pm 0.001 \times 10^{-12}$g/cell by the end. Given that mitochondrial biogenesis is occurring, it requires the synthesis of many important mitochondria genes. We also assessed the expression of COX1, COX2, and COX3 that are essential for oxidative phosphorylation and encoded in the mitochondria DNA and revealed their increase in protein expression over the course of the differentiation (Figure 6-10C).

6.2.6 Amino Acid Contribution to the Energy Metabolism During Hepatocyte Differentiation

In order to characterize the metabolic changes that cells undergo during hepatocyte differentiation, we also measured the amino acid concentration (Figure 6-11 and 6-12). All essential amino acids were consumed over the course of the differentiation (Figure 5A). However, the demand for amino acid was observed to lessen over the course of differentiation. Similarly, the metabolic demand for non-essential amino acid was also less over the course of the differentiation. Certain non-essential amino acids (e.g. alanine, glutamate and proline) were secreted by the cells. Alanine and glutamate have been frequently reported as the major amino acids that are produced by different types of stem cells (171).

Using the specific rates of the different metabolites, a metabolic flux analysis was performed and the average fluxes could be mapped onto the central metabolic pathways (Figure 6-14). Consistent with the specific rates obtained through experiments, the glycolytic flux ($J_{GLC\text{-}PYR}$) was twice as high in the first stage of differentiation compared to the last stage (Figure 6-13). Although, most of the fluxes within the TCA cycle were comparable throughout the differentiation but the first few pathways in the TCA cycle reveals a higher flux after the first stage of differentiation. The observed increase in TCA flux over the course of the differentiated could be attributed to the increase in ($J_{PYRC\text{-}PYR}$), responsible for the transport of pyruvate into the mitochondria after glycolysis which is consistent with our previous observation with an increase expression in the enzyme facilitating the transport, MPC. We also observe an increase in mala-aspartate shuttle fluxes suggesting that oxidative phosphorylation is more active in later stages of differentiation.

6.2.7 Influence of Glucose on Metabolism During Stem Cell Differentiation

From the previous sections, we observed that cells transition from a high to low glycolytic state over the course of the differentiation starting in the endodermal stage when cells begin to commit to the hepatic lineage. To show the importance of the metabolic switch, we used varying levels of glucose concentration to manipulate the metabolism state

of the cells during the differentiation. We observed that cells cultured in high glucose concentration at ~25mM consumed much larger quantities of glucose and secreted larger quantities of lactate (Figure 6-15). We compared differentiating hESCs at high glucose concentration with the conventional concentration at ~5mM and showed that the metabolic switch from high to low glycolytic state does not occur in the presence of high glucose concentration. Interestingly, the rate of glucose consumption and lactate production on a per cell basis was almost triple of what was observed in the high glucose differentiation compared to the conventional glucose concentration. Using qRT-PCR and albumin secretion assay, we showed that differentiation carried out at high glucose concentration were still capable of exhibiting hepatic phenotypes but to a lesser degree compared to a differentiation carried out a lower glucose levels (Figure 6-16).

6.3 Discussion

Although the energetic state of hESCs have been studied extensively, the mechanisms and observation behind the bioenergetic changes during hepatic differentiation are still not completely understood. The present study assessed the metabolism of hESCs during the differentiation towards the hepatic lineage and how it can influence the way we carry out future processes for stem cell culture. Based on our findings, we can conclusively state that OxPhos is activated during the hepatocyte differentiation of hESCs. Compared to other studies documenting the bioenergetic state of undifferentiated and differentiated cells, our study evaluates the bioenergetics demands and mitochondria network at different time points over the course of hepatocyte differentiation for hESCs.

In this study, we explored into the metabolic mechanisms and energetic requirements that are occurring over the course of the differentiation. With regards to glycolysis, cells in the first stage of differentiation adopt a high glycolytic state where high glucose consumption and high lactate production was observed. This phenomenon is typically referred to as the Warburg Effect and has been observed in highly proliferative cells such as cancer cells and stem cells for decades [insert citation]. Our study reveals that differentiating endodermal cells also retain the pre-existing high glycolytic state. It is evident that the highest level of glycolysis occurs during the beginning of the differentiation and begins to transition towards a low glycolytic state starting after committing to the hepatic lineage during the second stage of differentiation. This decrease

in glycolytic capability is accompanied by the transition towards a OxPhos metabolism. We begin to observe a gradual increase in oxygen consumption starting in the second differentiation stage and reaching peak levels of 2.3 x 10^{-10} mmol/cell/h by the end of differentiation for SCDHs. Many studies have attempted to measure the intrinsic oxygen consumption rate of different primary hepatocytes and reported of similar values between 3 x 10^{-10} – 3 x 10^{-9} mmol/cell/h (172-175).

It is widely accepted that differentiated and metabolically active cells such as primary hepatocytes have more efficient mitochondrial metabolism in favor of OxPhos. The plasticity in energy metabolism allows stem cells to change their metabolic phenotype to satisfy the different needs of self-renewal and its different lineage specification. Our findings suggest that when stem cells commit to the hepatic lineage there is a smaller dependence on glycolysis accompanied by an increase in oxygen consumption. One possible explanation for the increase in oxygen consumption could be an increase in mitochondria biogenesis. To assess the generation of mitochondria, we measured the mtDNA levels on a per cell basis and found significant increase over the course of the differentiation. We further seek to investigate whether mitochondria network formation was also occurring to promote the activation of OxPhos during our hepatic differentiation. Typically, immature mitochondria with less defined cristae are more common among cells that depend more on glycolysis (176). It has frequently been reported that mature mitochondria is crucial and required during stem cell differentiation for cells that require OxPhos especially in cardiac differentiation (177-179). In addition to the increase of mtDNA levels over the course of hepatic differentiation, the membrane potential of the mitochondria was assessed using a rhodamine dye which revealed an increase in mitochondria activity supporting our hypothesis that OxPhos has been activated. The maturity of mitochondria has been shown to play a vital role in the balance between pluripotency and cellular differentiation(180). In this study, we further demonstrated the maturity of our HLCs through the established network formation. Thus, we can conclude that the increase in OxPhos during hepatic differentiation is due not only to mitochondria generation but also the maturation of mitochondria as well.

The observed metabolic switch and increase in mitochondria biogenesis must also have a regulatory mechanism during differentiation. We used transcriptome data to look at glucose-related genes to give us insight into what possible regulators might be behind the changes in energy metabolism. Our data indicated that there may not be a master regulator behind the metabolic switch and decrease of glycolytic activity. Instead, we observed a number of different glycolytic enzymes showing decreased expression over the course of the differentiation. Specifically, we observed a decrease in the transporter (GLUT1 and GLUT3) responsible for providing intake of glucose to the cells. Interestingly, we also show that there is a decrease of expression in both predominant forms of hexokinase (HK1 and HK2). Several studies have shown the elevated glucose uptake from the overexpression of simply either GLUT or HK or both (181, 182). Although we observed a number of different glycolytic enzymes showing decreased expression in support of our metabolic switch, most metabolic enzymes in the TCA cycle are present at very high levels throughout the differentiation without any remarkable changes during differentiation. However, mitochondria pyruvate carrier (MPC) was expressed at very low levels until its induction during the second stage of differentiation corresponding to the increase in OxPhos activity. Recent works have shown that MPC is vital to hepatic gluconeogenesis and TCA cycle flux (183, 184). The loss of MPC activity was shown to impair the transport of pyruvate to the TCA cycle and ultimately diminishing the levels of metabolic intermediates found in the mitochondria. We confirmed the importance of MPC for HLCs generation by supplementing UK5099, a MPC inhibitor, and observed that cells were unable to differentiate past the second stage of the differentiation ultimately resulting in cell death (data not shown). This allowed us to conclude that the metabolic switch to OxPhos is essential to commitment to the hepatic lineage. Despite all the work conducted, the exact mechanism driving the switch to OxPhos to induce TCA flux has not yet been described.

In this study, we have described the bioenergetic needs of cells at different stages of hepatic differentiation and identified the metabolic switch that occurs which is crucial for making HLCs. Together, our findings provide important insights into the distinct regulation of bioenergetics during hepatic differentiation and the importance of the metabolic switch. At later stages of differentiation when hepatic cells begin to exhibit more

mature processes of lipid and carbohydrate metabolism, the demand for active mitochondria and TCA flux increases while the need for glycolysis diminishes. The work conducted here will be important as we move forward exploring new approaches to manipulate hESCs fate and lineage commitment to generating mature HLCs for new and important therapeutic opportunities.

6.4 Figures

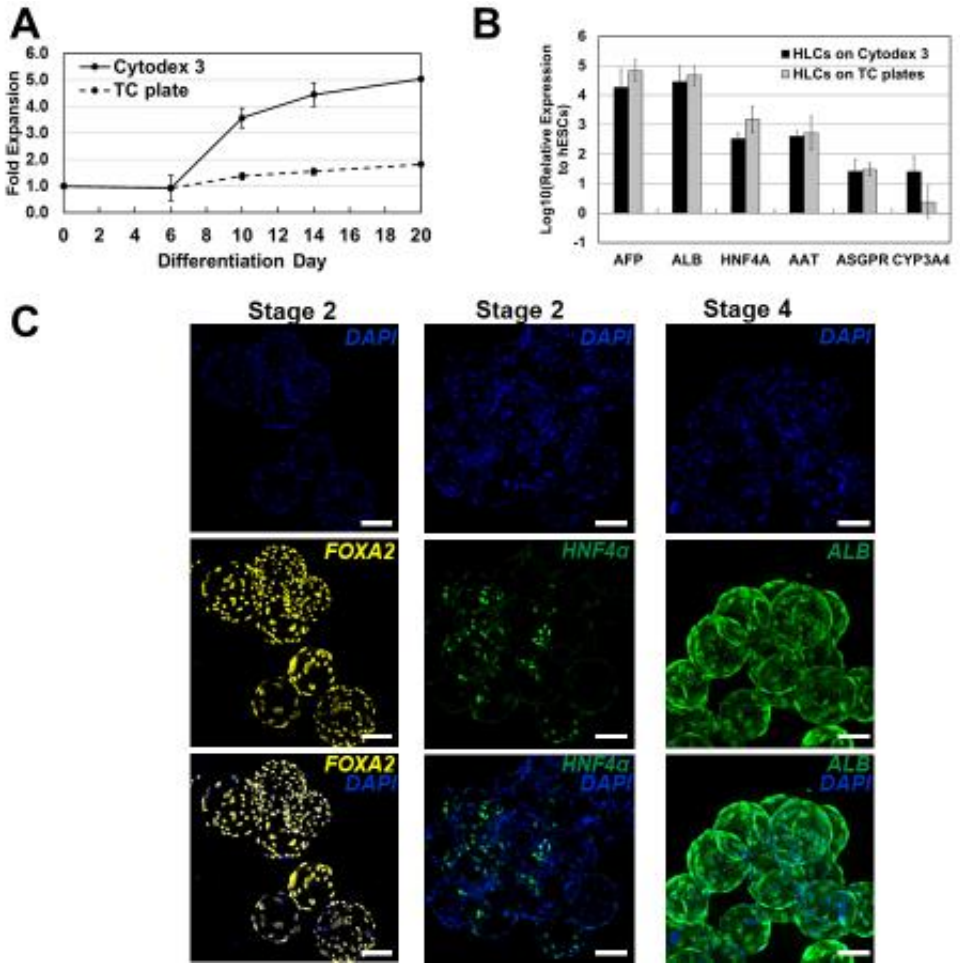


Figure 6-1. Characterization of HLC Differentiation. A) Cell count during hepatocyte differentiation on Cytodex 3 microcarriers (bold line) and on tissue culture (TC) plate (dashed line). Error bars are depicted as standard deviation of 3 replicates. B) Quantitative RT-PCR (qRT-PCR) of different hepatic markers of HLCs normalized to GAPDH compared to human embryonic stem cells (hESCs) for differentiation on Cytodex 3 and TC plates. Error bars are depicted as standard deviation of 3 replicates. C) Immunohistochemistry of endodermal, *FOXA2*, and hepatic (*HNF4α, ALB*) markers at different stages of differentiation on Cytodex 3.

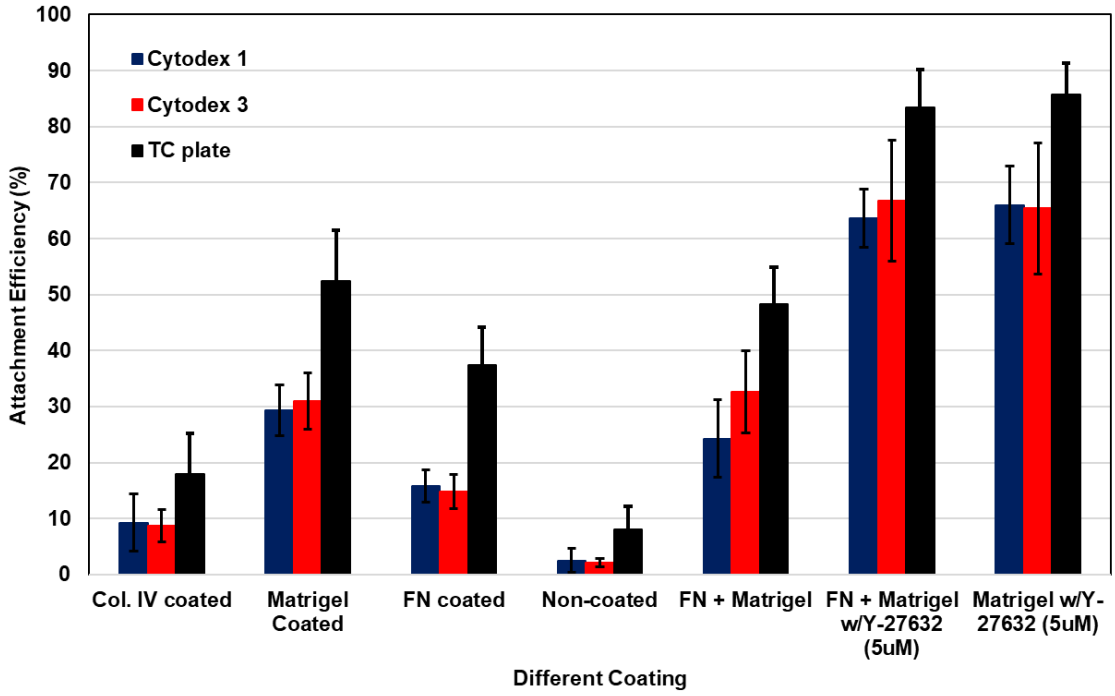


Figure 6-2. Attachment Efficiency on Microcarriers. Different microcarriers (cytodex 1 and 3) was examined for attachment efficiency compared to tissue culture plate (TC plate). Different coating using Matrigel/Collagen IV/ and Fibronectin (FN) was tested to give the best attachment efficiency.

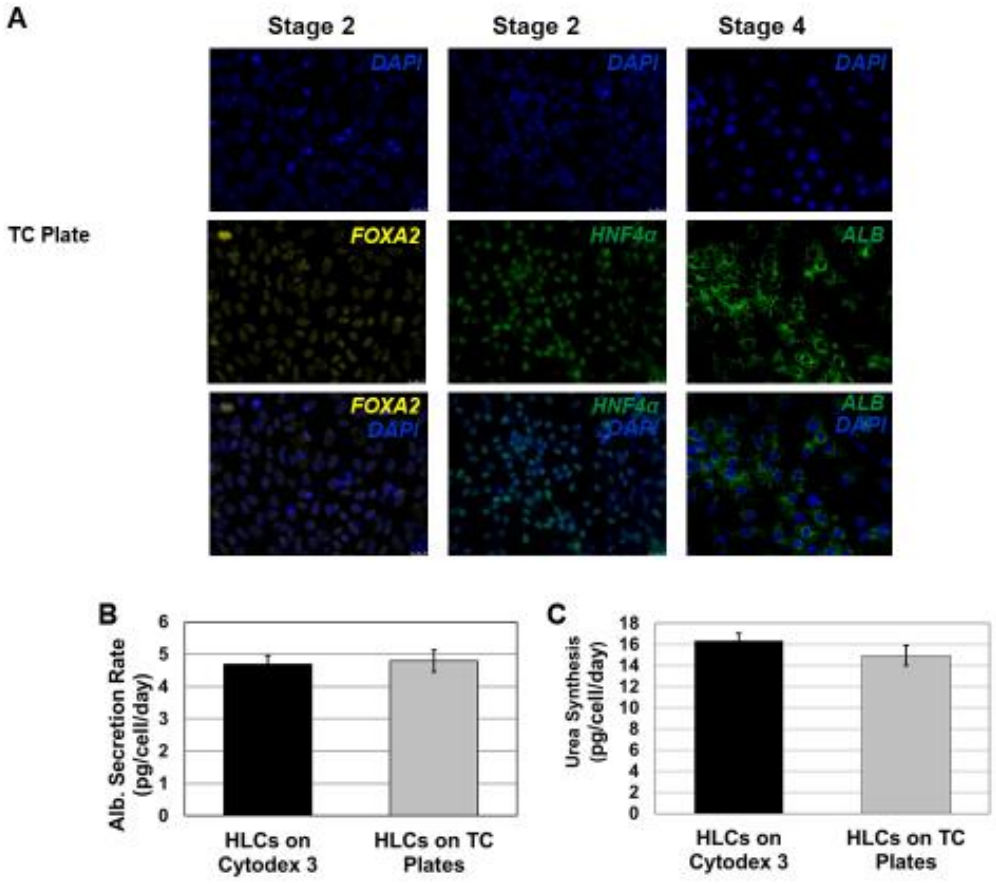


Figure 6-3. Functional Characterization of HLC Differentiation. A) Immunohistochemistry of endodermal, *FOXA2*, and hepatic (*HNF4α, ALB*) markers at different stages of differentiation on Cytodex 3 and TC plate. Scale bar is 50um. B and C) Albumin and urea synthesis for HLCs differentiated on TC plate.

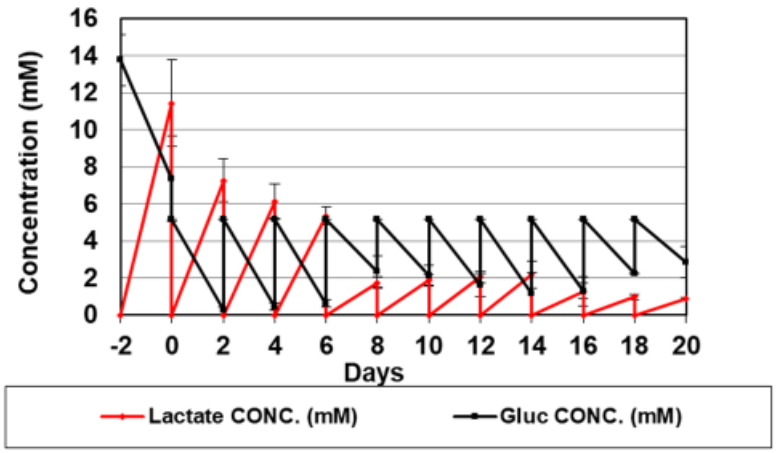


Figure 6-4. Glucose and Lactate Concentration Profile over Course of Differentiation.

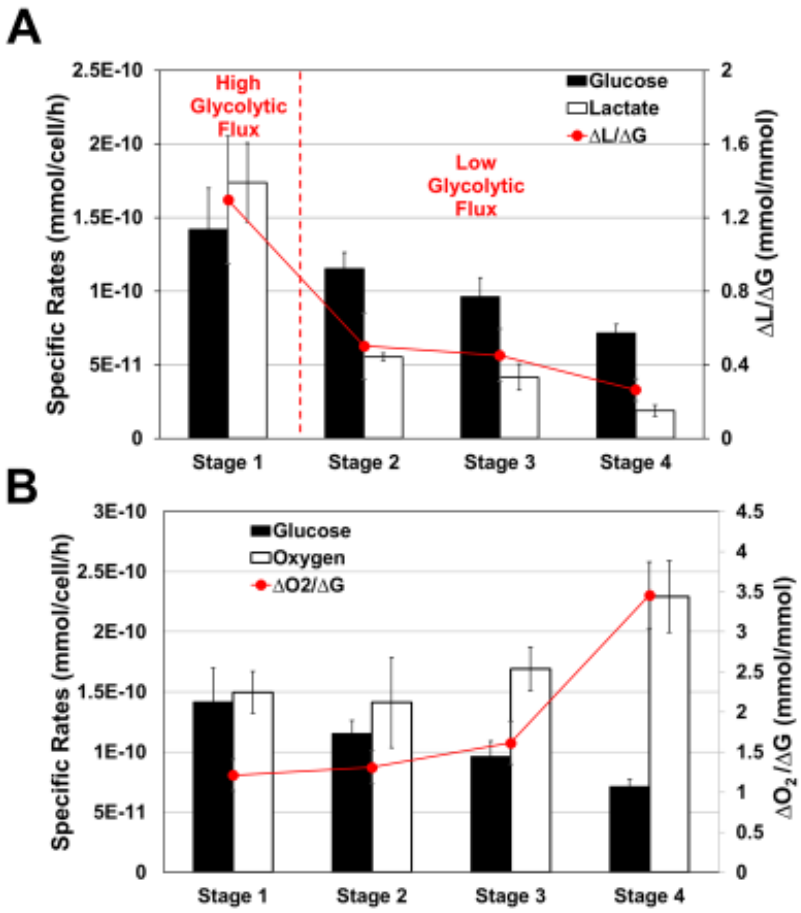


Figure 6-5. Metabolism Profiling of Cells During HLC Differentiation. A) Glucose consumption and lactate production rate for cells during different stages of differentiation. Stoichiometric ratio of lactate:glucose was plotted in a solid red line. B) Oxygen consumption was plotted along with glucose consumption for different stages. Stoichiometric ratio of oxygen:glucose was plotted in a solid red line.

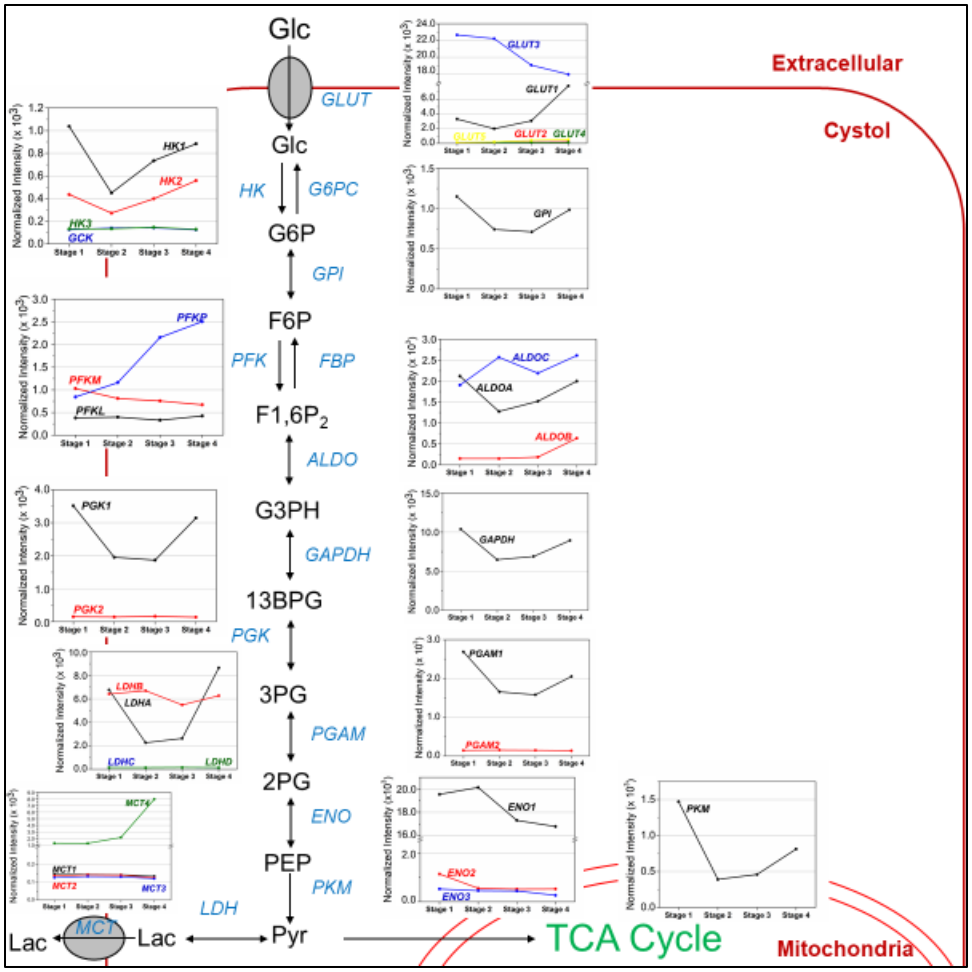


Figure 6-6. Microarray Data on Different Glycolytic Isozymes During Differentiation. Different glycolytic enzymes were plotted along with its different isozymes. The data was linear normalized and each color represents an isozyme.

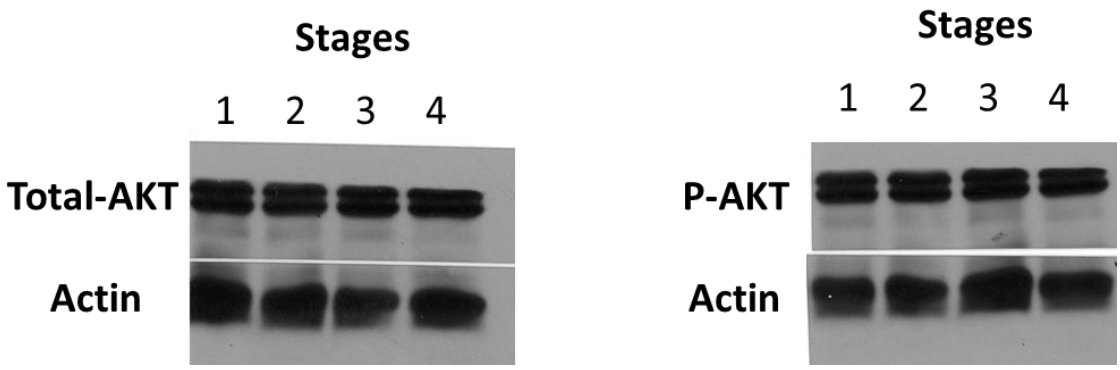


Figure 6-7 Western Blot of Total AKT and Phosphorylated AKT during Differentiation.

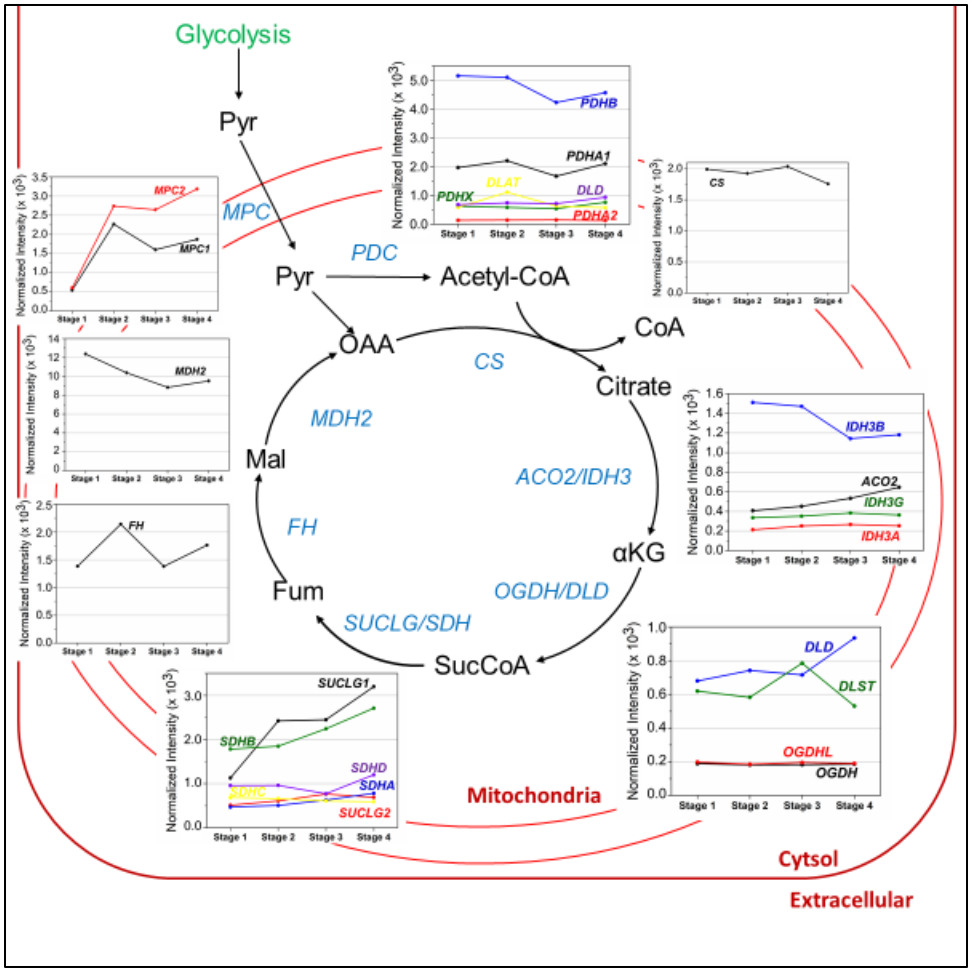


Figure 6-8. Microarray Data on Different TCA Cycle Isozymes During Differentiation. Different TCA cycle enzymes were plotted along with its different isozymes. The data was linear normalized and each color represents an isozyme.

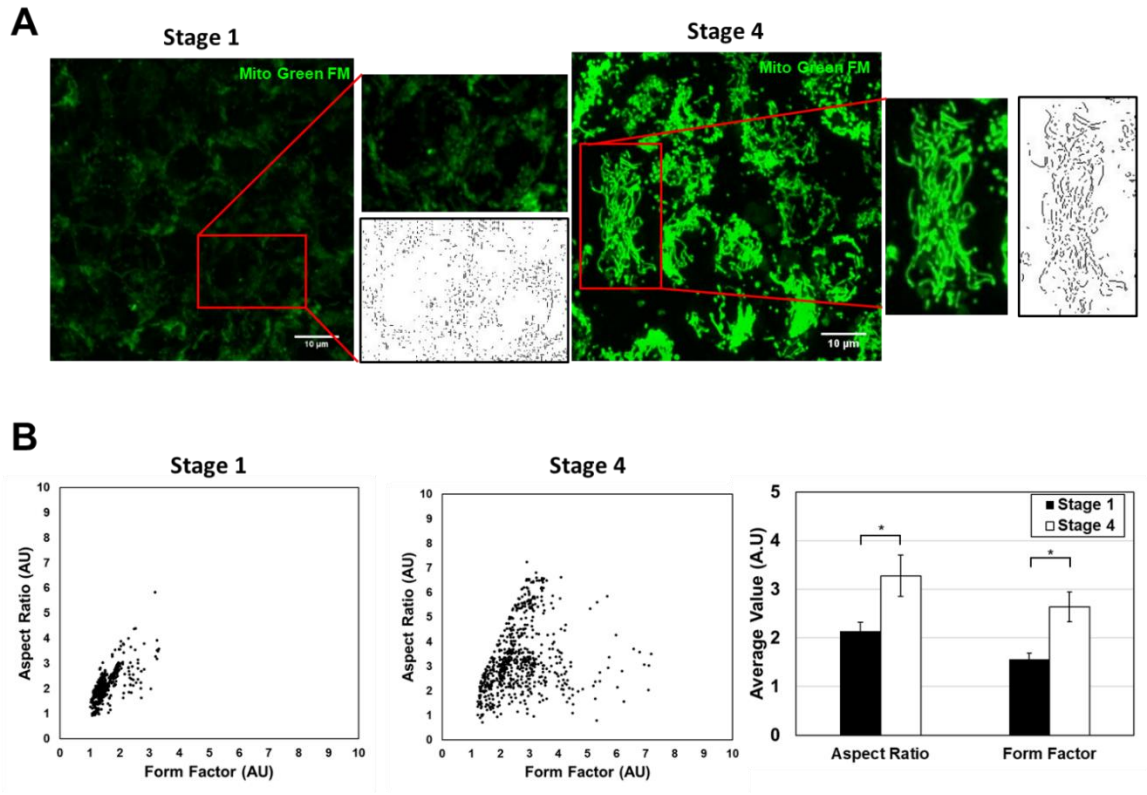


Figure 6-9. Assessment of Mitochondria Maturation During Differentiation. A) Mitotracker Green FM Staining on Stage 1 and Stage 4 cells. B) Form Factor and Aspect Ratio quantification on Stage 1 and Stage 4 cells.

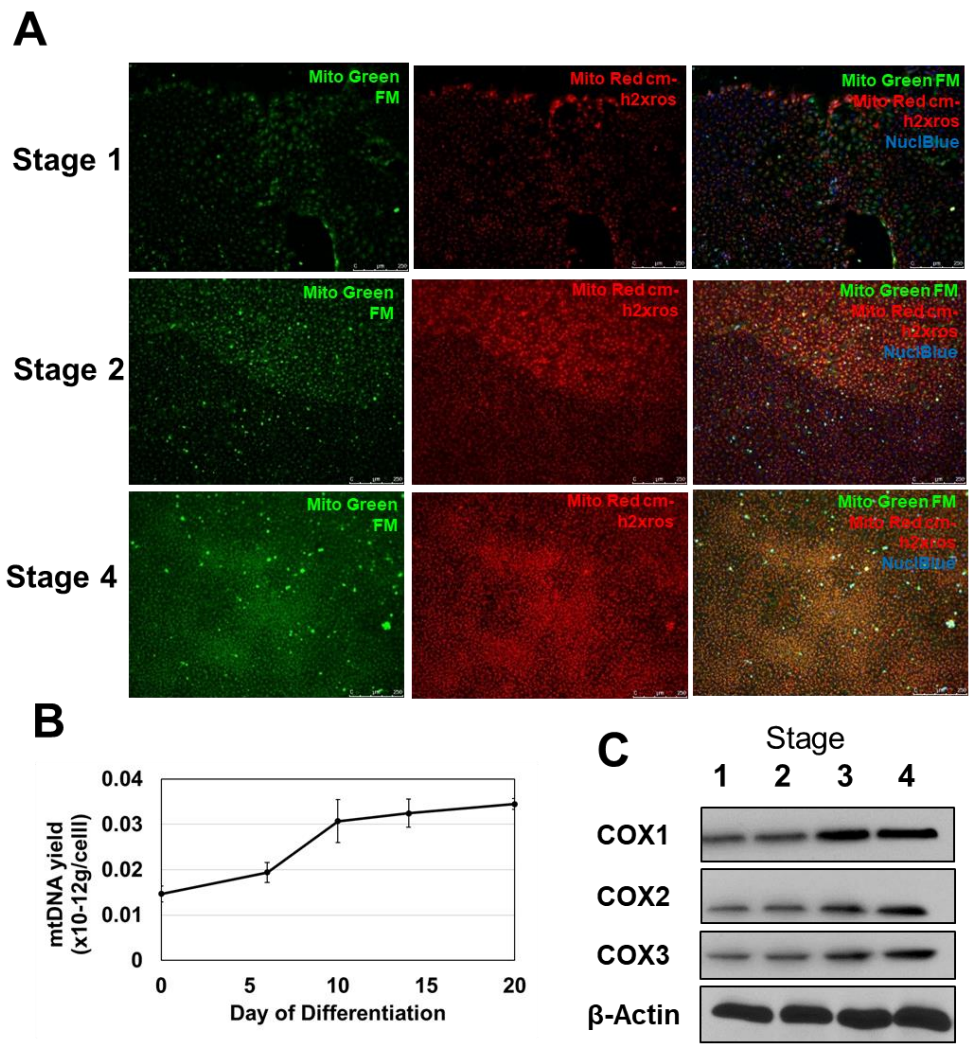


Figure 6-10. Mitochondria Activity and Biogenesis During Differentiation. A) Mitotracker Green FM and CM h2x-ros staining during differentiation. B) Quantification of mtDNA on cells at different days of differentiation on a per cell basis. C) Western Blot showing protein presence of different COX proteins during differentiation.

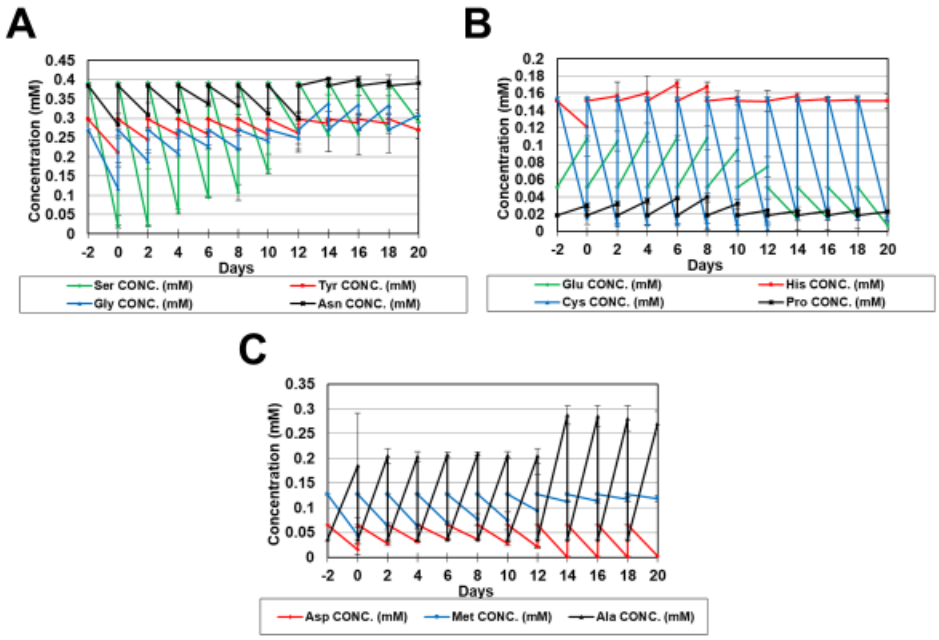


Figure 6-11. Non-essential Amino Acid Concentration Profile Over the Course of Hepatocyte Differentiation. A) Serine, Tyrosine, Glycine, and Asparagine. B) Glutamate,histidine, cysteine, proline. C) Aspartate, methionine, alanine

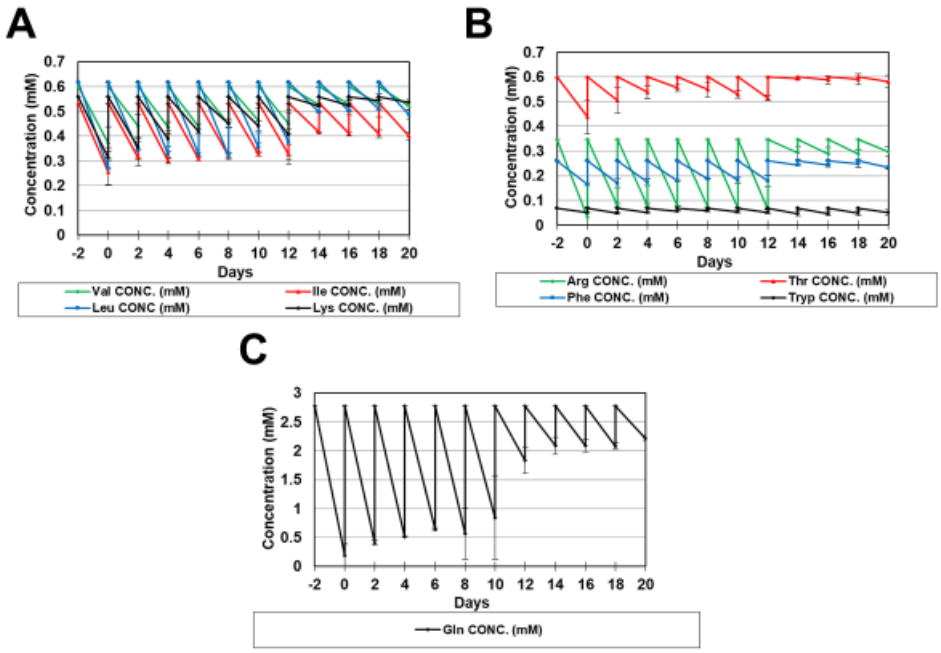


Figure 6-12. Essential Amino Acid Concentration Profile Over the Course of Differentiation. A) Valine, isoleucine, leucine, lysine. B) Arginine, Threonine, Phenylalanine, Tryptophan. C) Glutamine

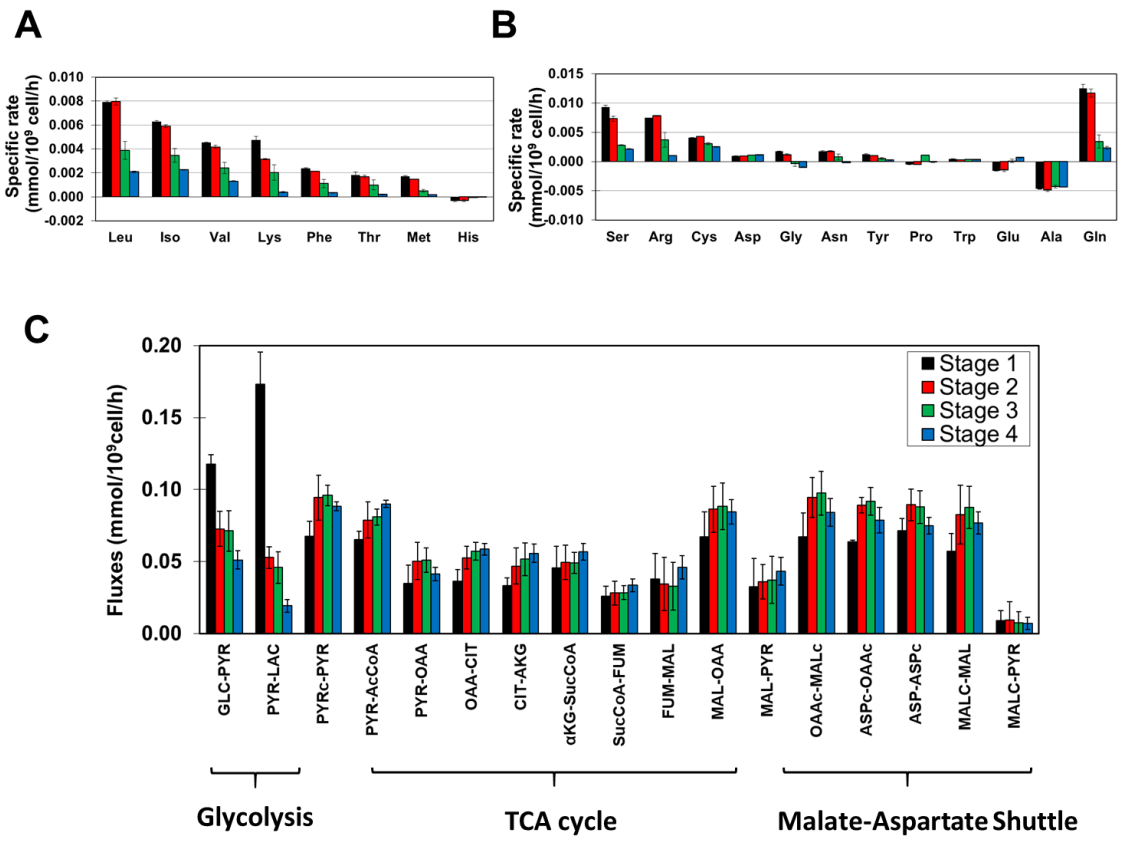


Figure 6-13. Metabolic Fluxes and Consumption Profiles. A) Consumption rate of different essential amino acids. B) Consumption/production rate of different non-essential amino acids. C) Metabolic flux (mmol/10^9cells/h) for different metabolic pathways

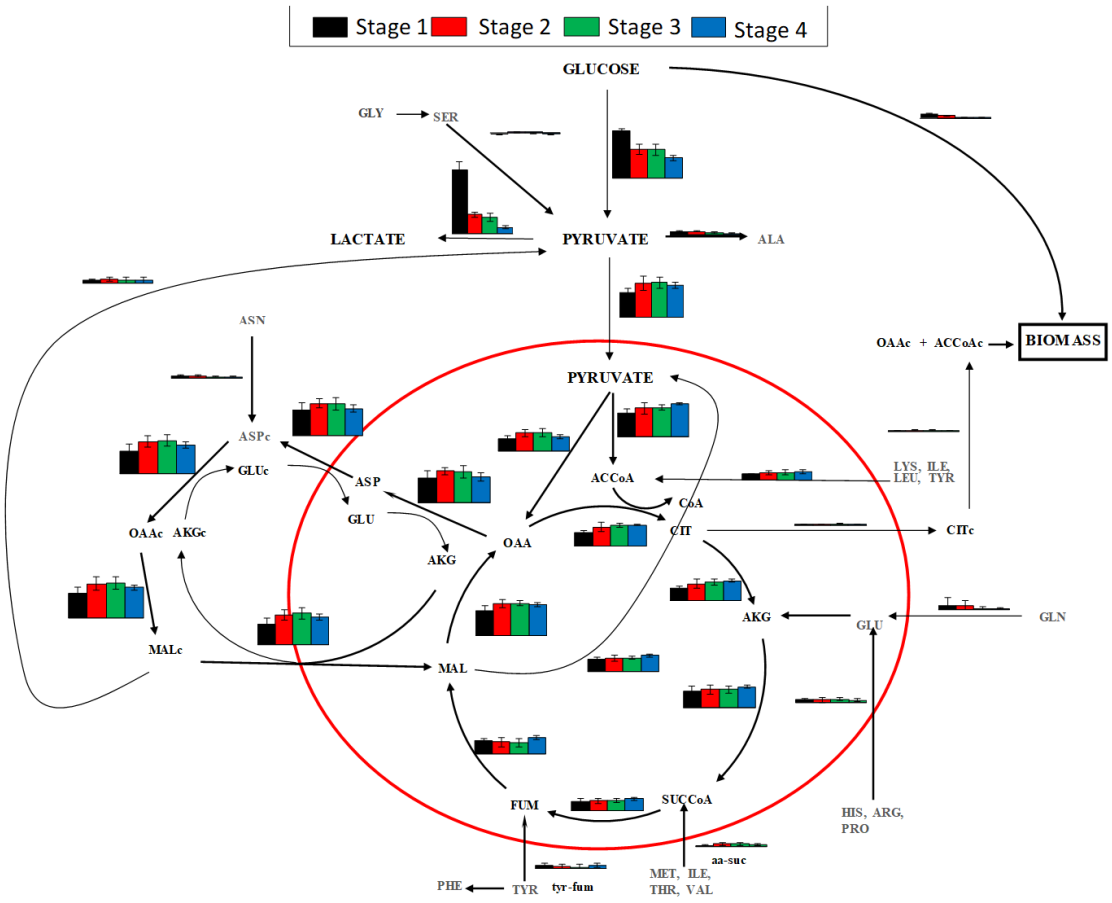


Figure 6-14. Metabolic Fluxes Chart. Metabolic flux (mmol/10^9cells/h) mapped onto metabolic pathway chart with the same scale for every chart.

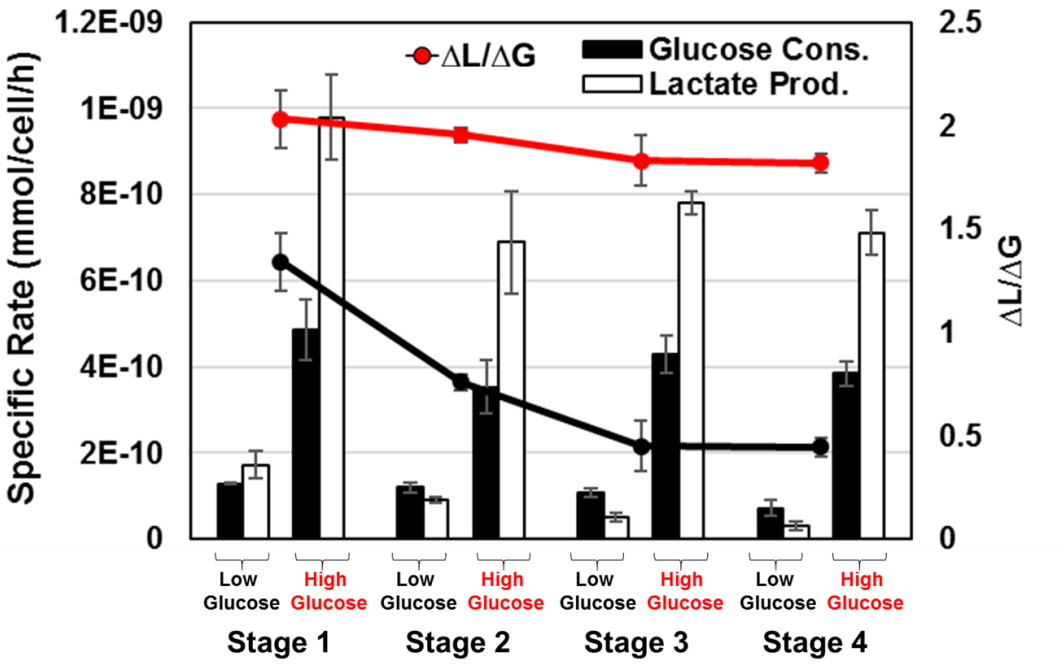


Figure 6-15. Effect of Glucose Concentration on Glucose and Lactate Profile Over the Course of the Differentiation. Glucose consumption (black) and lactate production (white) were measured over the course of the differentiation using low glucose (1g/L) and high glucose (4.5g/L) differentiation conditions. Stoichiometric ratio of lactate: glucose (line graph) were plotted to illustrate the glycolytic activity during differentiation.

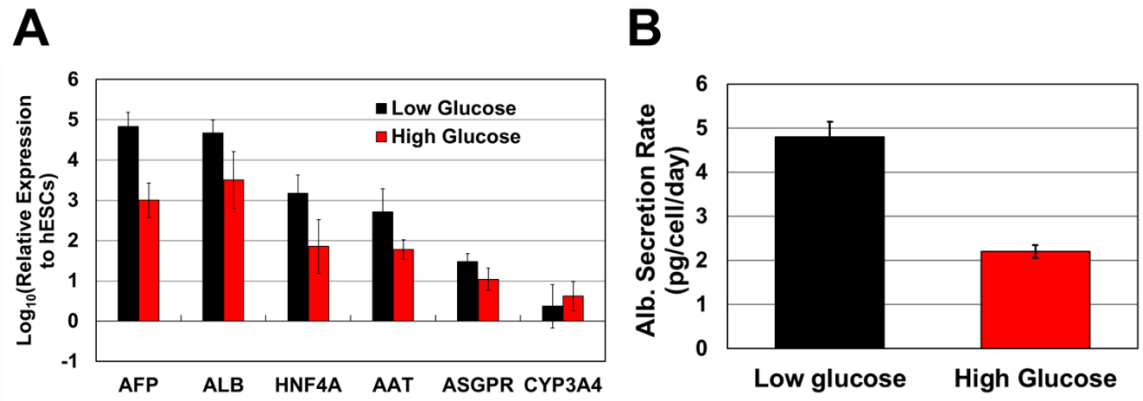


Figure 6-16. Effect of Glucose Concentration on Differentiation of HLCs. A) Transcript levels of different hepatic markers on HLCs using high and low glucose for differentiation B) Albumin secretion rate for differentiation carried out in high and low glucose levels

Chapter 7. Conclusions and Future Work

In the last few decades, large interest and research efforts have gone into harnessing the full potential of stem cells. With the capability to self-renew while being able to differentiate into any specialized cell types in response to different signaling cues, stem cells can become the next revolution of medicine within the next few decades. These intrinsic properties of stem cells provide great promise in using them to differentiate into hepatocyte for different applications such as drug screenings, bioartificial liver devices and tissue engineering. With the not too recent discovery of iPSCs, even more attention and funding have poured into using these particular type of stem cells to treating different illnesses without the potential of an immune response. Currently, there are already over 600 clinical trials being conducted using stem cells and 30 of them are focused on treating different liver disorders and diseases. Hopefully, within the next few decades, the reality of using stem cells will have been achieved. For this potential to be realized however, advancements must be made in all phases of translating the laboratory practices we have in place to a more efficient and scalable bioprocess. Therefore, to be able to completely reach the full potential of stem cells, we must first be able to establish a process to be able to consistently and safely produce large quantities of high quality hepatocytes.

In this study, we utilized a systems approach to address different aspects of clinical translation of stem cell research. These key objectives were divided into a) increasing product quantity (Chapter 4) b) enhancing product quality (Chapter 5) from stem cell differentiation and c) understanding the metabolic effects to develop a robust and scalable bioprocess for large scale differentiation (Chapter 6).

7.1 Process Improvement

In chapter 4, we showed that our current differentiation protocol can be optimized further to generate larger quantities of HLCs. We utilized the intrinsic properties of endodermal cells to provide large quantities without the need to culture on xenogeneic feeder cells or cell sorting that previous groups have shown. These cells were given the signaling cues of BMP4 and FGF2 with extra surface area for growth and showed the capability to expand up to 15-fold. These factors were chosen based on the natural

mesodermal signaling cues that endodermal cells are exposed to during *in-vivo* liver development. This observation of cell expansion showed a simultaneous 1000-fold increase in different hepatic markers and 1000-fold decrease in endodermal and pluripotent markers indicating the ability of cells to simultaneously expand and commit to the hepatic lineage similar to in-vivo embryogenesis. To determine the efficiency of this process at a single cell-resolution, we utilized mass cytometry to probe the co-expression of different endodermal and hepatic markers at different stages of the conventional differentiation and the modified differentiation. We were, to the best of our knowledge, the first to use mass cytometry for hepatic stem cell differentiation. Using this technique, we showed that endodermal and hepatic genes were repressed and expressed simultaneously for a successful outcome. Mass cytometry allows us to do a more thorough characterization of our cells by being able to look at different markers without running into spectra overlap faced in flow cytometry. Through mass cytometry, functional analysis, and transcriptome analysis, we showed that the end stage HLCs obtained from the modified version were comparable and in some instances better than the conventional method of differentiation. We hypothesized that by providing extra surface area during the process to allow endodermal cells to proliferate it more accurately mimics what occurs *in-vivo* to give better HLCs.

7.2 Quality Enhancement

In chapter 5, we discussed the functional aspect of the current efforts of hepatic stem cell differentiation and the current drawback faced in the field. Although our HLCs exhibited certain characteristics of primary hepatocytes, the levels of function and gene expression of HLCs still fall short compared to primary hepatocytes to be used in clinical settings. Through surveying recent literature and examining our own HLCs, it appears that all HLCs generated have the same level of immaturity and characteristics of fetal hepatocytes regardless of the differentiation protocol or optimization strategies used. Thus, we employed using a meta-analysis approach to understand the genetic roadblocks that are preventing stem cells from reaching the functional maturity of primary hepatocytes. Hence, it seemed logical to mimic embryonic liver development as stem cells were harvested during the blastocyst stage of development. However, transcriptome data of human *in-vivo*

development is very hard to obtain due to ethical reasons. As a result, we assumed that certain aspects of mammalian development are conserved between mouse and human development. Therefore, we succeeded in combining human *in-vitro* differentiation transcriptome data with mouse *in-vivo* development transcriptome data to give us insight into strategies for enhancing maturation.

We first obtained transcriptome data at different stages of our *in-vitro* human differentiation. We first examined certain families of mature hepatic genes in our cells that are known to be important for hepatocyte phenotype and function. These include the liver-enriched hepatocyte nuclear factors (HNFs), Cytochrome P450 enzymes, and gluconeogenesis enzymes. We observed most genes were expressed at moderate levels but not at levels that should represent primary hepatocytes. In some instances, certain mature hepatic markers are not expressed at all such as FBP1 or GCK. To reinforce our assessment of HLCs, we combined different human *in-vitro* hepatic differentiation from different groups in the public databases. These datasets, however, originated from different assay platforms and sources. As a result, we needed to treat the data with a ComBat algorithm in order to get rid of certain batch effects that arise from compiling different datasets. This allowed us to perform a meta-analysis by compiling many different datasets across different studies in order to obtain a more comprehensive and conclusive finding. Using this same approach, we also incorporated *in-vivo* developmental data where we saw that all *in-vitro* differentiation processes fall short of primary hepatocytes. But using developmental data, we were able to pinpoint the exact developmental stage that our HLCs are stuck at and align *in-vivo* and *in-vitro* data on the same scale for direct comparison. By creating a uniform "developmental time", we can now go into each gene for each process and observe the dynamics of the different genes. To our surprise, we observed a number of discordant genes that behave differently in our human *in-vitro* differentiation process compared to *in-vivo* development. These dynamically identified differentially expressed genes can provide us a way to further enhance the maturation of HLCs through certain genetic intervention or strategies to alter their expression dynamic.

The conclusion of this finding allowed us to develop a unified "developmental time" scale to compare different differentiation stages with its corresponding embryo

developmental stages. This development allowed us to conclude that most HLCs, including our own, are stuck at around E14.5 of development suggesting that 5 more developmental days of mouse signaling cues are needed to promote further maturation before achieving primary hepatocyte status. This analysis uncovered a pivotal gene set of about 200 genes that can now allow us to narrow down the number of targets we can go in with to genetic intervene given the recent discovery of CRISPR-CAS9 which is beyond the scope of this study. However, we believe this uncovering and analysis will be key in moving the field forward and can also be applied to other applications besides hepatic stem cell differentiation given that enough datasets are available.

7.3 Process Development

As discussed previously, we enhanced our conventional differentiation process to improve cell quantity and looked into developing strategies for enhancing the cell quality. One of the last aspects that we attempted to address in this study was to develop a process that would be more scalable. We briefly explored into using different microcarriers and different coatings to develop a scalable manner of expanding and differentiating stem cells to hepatocytes. Through the development of this process, we also looked into how the microenvironment, specifically the presence of glucose, can affect the quality of hepatocytes we obtain and how that might affect our process.

From our previous transcriptome analysis and other reports, we know that gluconeogenesis is not present in our differentiation process but we did see a decrease in glucose consumption and glycolysis activity. We succeeded in measuring and determining the different metabolic demand of our cells to be able to characterize our cells. We observed that cells have a smaller demand for glucose and amino acids throughout the course of the differentiation. Without the need to proliferate and expand, the need for glucose and other nutrients are smaller for HLCs compared to their stem cell counterpart. We did observe a metabolic switch from glycolysis to oxidative phosphorylation suggesting to us that our HLCs do change their metabolic profile to more represent primary hepatocytes. However, we noticed that gluconeogenesis does not occur as a lot of the enzymes involved in gluconeogenesis is not expressed. Furthermore, we demonstrated that by using different concentrations of glucose, we can obtain different qualities of HLCs. Similar to other

mammalian cells that have been cultured, we observed that high presence of glucose can affect the glucose consumption and hence the glycolytic activity of our cells during the differentiation process. As a result, we found that differentiation carried out with lower levels of glucose gave more optimal differentiation of HLCs with higher expression of hepatic markers. We discovered that although glucose consumption and a number of different glycolytic enzymes were repressed over the course of the differentiation, the TCA activity and flux seemed to be going up over the course of the differentiation. From our metabolic flux analysis, we revealed that the gene, MPC, is responsible for the influx of carbons into the TCA cycle. Although, there is lower levels of glucose consumption, the sudden expression in MPC allows cells to more effectively channel pyruvate into the TCA cycle for mitochondria respiration reflected by the increase in oxygen consumption over the course of the differentiation. This information will allow us to fully define the metabolic demands of our cells at different stages of the differentiation to prevent overfeeding of certain nutrients where we showed glucose can be detrimental.

7.4 Future Work

This thesis work provides a thorough understanding of the current status of hepatic stem cell differentiation and the development of a process to translate laboratory stem cell practices into a stem cell technology. The availability of a renewable source of hepatocytes will have profound impact not only in the clinical settings but also in the industrial settings. HLCs have already been shown to be effective in the industrial setting as a platform for drug toxicity screenings. By relying on human HLCs, clinicians expect a more accurate screening using human cells that can represent human physiology instead of relying on animal models. However, in the future, cellular therapy and all the different applications will require high-quality, functionally mature hepatocytes in large quantities.

With the discovery of iPSCs, the discussion of personalized medicine for individuals or selective populations have transpired and picked up steam. Current work have already shown the capability of generating HLCs from ESCs or iPSCs using the same differentiation process. However, as we discussed in this thesis, the current field needs to address the ability to generate HLCs in a robust manner representative of primary hepatocytes.

The lack of functional maturity in HLCs were addressed in Chapter 5 where a thorough transcriptome analysis was conducted to show that there is a universal gap that is faced by all research groups regardless of differentiation strategy. This functional immaturity is a big concern and represents one of the biggest bottleneck that the field of stem cells is facing. In our study, we conducted a meta-analysis to identify certain genes and their regulators that might be key in enhancing functional maturity of HLCs. Additionally, it was revealed in this study that HLCs are blocked at a developmental stage corresponding to E14.5 in mouse development. Typically, HLCs are transplanted into animal models to show rescue of liver failure or restoration of liver function. However, we believe that HLCs are not at the maturity level to show any real progress by transplanting into adult livers of different models. We propose an alternative strategy to transplant HLCs into the embryonic livers of mice at E14. This approach will rely on the environmental cues of an E14 mice to further enhance the maturity of HLCs. By using a comparative transcriptome and proteomic analysis of both cells transplanted and without transplanted, we can then begin to understand the critical developmental cues that we are missing in our *in-vitro* systems.

Lastly, a final consideration that must be addressed for transplantation and different applications is determining what process can be robust enough to generate large quantities of HLCs. We showed that it is capable of differentiating HLCs in defined conditions without the use of serum that are not only suitable for laboratory practices but also for manufacturing and clinical applications. For most applications using adherent cells, microcarriers is a common matrix to support cell growth and adherence especially in the manufacturing of viral vaccines in large scale. We were able to show that microcarriers can also support stem cell growth and differentiation for facilitating large quantities of HLCs. However, more test will need to be conducted to determine the effect of mechanical stress on cells in large tank reactors. But we showed that using gene expression analysis we identified metabolic changes ESCs undergo as they transition towards HLCs. Using this knowledge, we identified that cells have different metabolic demands over the course of the differentiation and that certain overfeeding can hinder the differentiation process. For safety reasons and process development, we established a robust process capable of

differentiating endodermal cells on microcarriers in a scalable manner that can be utilized in large tank bioreactors.

With the recent advancements in developmental biology, it is an exciting time in the stem cell field as we continue to understand the control of cell fate. However, hopefully stem cells, in the next decade, will be key in realizing the potential impact of using stem cells as a technology in different clinical and industrial applications. This will require the collaborative efforts of researchers from academia and industry from different fields of study in order to establish this technology. Hopefully as we demonstrated in this study, the innovations and discovery being put into the area of stem cells will one day make regenerative medicine a reality in the not too distant future.

Chapter 8. References

1. Annual Data Report of the US Organ Procurement and Transplantation Network. Preface. Am J Transplant 2014;14 Suppl 1:5-7.

2. Hayflick L. The Limited in Vitro Lifetime of Human Diploid Cell Strains. Exp Cell Res 1965;37:614-636.

3. Morrison SJ, Prowse KR, Ho P, Weissman IL. Telomerase activity in hematopoietic cells is associated with self-renewal potential. Immunity 1996;5:207-216.

4. Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PL, Coviello GM, et al. Specific association of human telomerase activity with immortal cells and cancer. Science 1994;266:2011-2015.

5. Hiyama E, Hiyama K. Telomere and telomerase in stem cells. Br J Cancer 2007;96:1020-1024.

6. Ohtsuka S, Dalton S. Molecular and biological properties of pluripotent embryonic stem cells. Gene Ther 2008;15:74-81.

7. Ho L, Ronan JL, Wu J, Staahl BT, Chen L, Kuo A, Lessard J, et al. An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. Proc Natl Acad Sci U S A 2009;106:5181-5186.

8. Xi R, Xie T. Stem cell self-renewal controlled by chromatin remodeling factors. Science 2005;310:1487-1489.

9. Onder TT, Kara N, Cherry A, Sinha AU, Zhu N, Bernt KM, Cahan P, et al. Chromatin-modifying enzymes as modulators of reprogramming. Nature 2012;483:598-602.

10. Shi Y, Desponts C, Do JT, Hahm HS, Scholer HR, Ding S. Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. Cell Stem Cell 2008;3:568-574.

11. Huangfu D, Osafune K, Maehr R, Guo W, Eijkelenboom A, Chen S, Muhlestein W, et al. Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat Biotechnol 2008;26:1269-1275.

12. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007;131:861-872.

13. Jaenisch R, Young R. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell 2008;132:567-582.

14. Duncan AW, Dorrell C, Grompe M. Stem cells and liver regeneration. Gastroenterology 2009;137:466-481.

15. Lyngbaek S, Schneider M, Hansen JL, Sheikh SP. Cardiac regeneration by resident stem and progenitor cells in the adult heart. Basic Res Cardiol 2007;102:101-114.

16. Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A 1981;78:7634-7638.

17. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. Embryonic stem cell lines derived from human blastocysts. Science 1998;282:1145-1147.

18. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, et al. Induced pluripotent stem cell lines derived from human somatic cells. Science 2007;318:1917-1920.

19. Amit M, Margulets V, Segev H, Shariki K, Laevsky I, Coleman R, Itskovitz-Eldor J. Human feeder layers for human embryonic stem cells. Biol Reprod 2003;68:2150-2156.

20. Amit M, Shariki C, Margulets V, Itskovitz-Eldor J. Feeder layer- and serum-free culture of human embryonic stem cells. Biol Reprod 2004;70:837-845.

21. Chase LG, Firpo MT. Development of serum-free culture systems for human embryonic stem cells. Curr Opin Chem Biol 2007;11:367-372.

22. Zhao T, Zhang ZN, Rong Z, Xu Y. Immunogenicity of induced pluripotent stem cells. Nature 2011;474:212-215.

23. Nussbaum J, Minami E, Laflamme MA, Virag JA, Ware CB, Masino A, Muskheli V, et al. Transplantation of undifferentiated murine embryonic stem cells in the heart: teratoma formation and immune response. FASEB J 2007;21:1345-1357.

24. Le Blanc K, Tammik C, Rosendahl K, Zetterberg E, Ringden O. HLA expression and immunologic properties of differentiated and undifferentiated mesenchymal stem cells. Exp Hematol 2003;31:890-896.

25. Li L, Baroja ML, Majumdar A, Chadwick K, Rouleau A, Gallacher L, Ferber I, et al. Human embryonic stem cells possess immune-privileged properties. Stem Cells 2004;22:448-456.

26. Hanna J, Wernig M, Markoulaki S, Sun CW, Meissner A, Cassady JP, Beard C, et al. Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. Science 2007;318:1920-1923.

27. Overturf K, al-Dhalimy M, Ou CN, Finegold M, Grompe M. Serial transplantation reveals the stem-cell-like regenerative potential of adult mouse hepatocytes. Am J Pathol 1997;151:1273-1280.

28. Desmet V, Roskams T, Van Eyken P. Ductular reaction in the liver. Pathol Res Pract 1995;191:513-524.

29. Kuwahara R, Kofman AV, Landis CS, Swenson ES, Barendswaard E, Theise ND. The hepatic stem cell niche: identification by label-retaining cell assay. Hepatology 2008;47:1994-2002.

30. Schmelzer E, Zhang L, Bruce A, Wauthier E, Ludlow J, Yao HL, Moss N, et al. Human hepatic stem cells from fetal and postnatal donors. J Exp Med 2007;204:1973-1987.

31. Tanaka M, Miyajima A. Identification and isolation of adult liver stem/progenitor cells. Methods Mol Biol 2012;826:25-32.

32. Dolle L, Best J, Empsen C, Mei J, Van Rossen E, Roelandt P, Snykers S, et al. Successful isolation of liver progenitor cells by aldehyde dehydrogenase activity in naive mice. Hepatology 2012;55:540-552.

33. Schmelzer E, Wauthier E, Reid LM. The phenotypes of pluripotent human hepatic progenitors. Stem Cells 2006;24:1852-1858.

34. Yoon SM, Gerasimidou D, Kuwahara R, Hytiroglou P, Yoo JE, Park YN, Theise ND. Epithelial cell adhesion molecule (EpCAM) marks hepatocytes newly derived from stem/progenitor cells in humans. Hepatology 2011;53:964-973.

35. Yamashita T, Forgues M, Wang W, Kim JW, Ye Q, Jia H, Budhu A, et al. EpCAM and alpha-fetoprotein expression defines novel prognostic subtypes of hepatocellular carcinoma. Cancer Res 2008;68:1451-1461.

36. Yamashita T, Budhu A, Forgues M, Wang XW. Activation of hepatic stem cell marker EpCAM by Wnt-beta-catenin signaling in hepatocellular carcinoma. Cancer Res 2007;67:10831-10839.

37. Turner R, Lozoya O, Wang Y, Cardinale V, Gaudio E, Alpini G, Mendel G, et al. Human hepatic stem cell and maturational liver lineage biology. Hepatology 2011;53:1035-1045.

38. Tanaka M, Itoh T, Tanimizu N, Miyajima A. Liver stem/progenitor cells: their characteristics and regulatory mechanisms. J Biochem 2011;149:231-239.

39. Kung JW, Forbes SJ. Stem cells and liver repair. Curr Opin Biotechnol 2009;20:568-574.

40. Zhang L, Theise N, Chua M, Reid LM. The stem cell niche of human livers: symmetry between development and regeneration. Hepatology 2008;48:1598-1607.

41. Wang Y, Cui CB, Yamauchi M, Miguez P, Roach M, Malavarca R, Costello MJ, et al. Lineage restriction of human hepatic stem cells to mature fates is made efficient by tissue-specific biomatrix scaffolds. Hepatology 2011;53:293-305.

42. D'Amour KA, Agulnick AD, Eliazer S, Kelly OG, Kroon E, Baetge EE. Efficient differentiation of human embryonic stem cells to definitive endoderm. Nat Biotechnol 2005;23:1534-1541.

43. Agius E, Oelgeschlager M, Wessely O, Kemp C, De Robertis EM. Endodermal Nodal-related signals and mesoderm induction in Xenopus. Development 2000;127:1173-1183.

44. Duboc V, Lapraz F, Saudemont A, Bessodes N, Mekpoh F, Haillot E, Quirin M, et al. Nodal and BMP2/4 pattern the mesoderm and endoderm during development of the sea urchin embryo. Development 2010;137:223-235.

45. Gadue P, Huber TL, Paddison PJ, Keller GM. Wnt and TGF-beta signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells. Proc Natl Acad Sci U S A 2006;103:16806-16811.

46. Lowe LA, Yamada S, Kuehn MR. Genetic dissection of nodal function in patterning the mouse embryo. Development 2001;128:1831-1843.

47. Bakre MM, Hoi A, Mong JC, Koh YY, Wong KY, Stanton LW. Generation of multipotential mesendodermal progenitors from mouse embryonic stem cells via sustained Wnt pathway activation. J Biol Chem 2007;282:31703-31712.

48. Cheng X, Ying L, Lu L, Galvao AM, Mills JA, Lin HC, Kotton DN, et al. Self-renewing endodermal progenitor lines generated from human pluripotent stem cells. Cell Stem Cell 2012;10:371-384.

49. Kamiya A, Kinoshita T, Miyajima A. Oncostatin M and hepatocyte growth factor induce hepatic maturation via distinct signaling pathways. FEBS Lett 2001;492:90-94.

50. Sekiya S, Suzuki A. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. Nature 2011;475:390-393.

51. Huang P, He Z, Ji S, Sun H, Xiang D, Liu C, Hu Y, et al. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. Nature 2011;475:386-389.

52. Wurm FM. Production of recombinant protein therapeutics in cultivated mammalian cells. Nat Biotechnol 2004;22:1393-1398.

53. Serra M, Brito C, Correia C, Alves PM. Process engineering of human pluripotent stem cells for clinical application. Trends Biotechnol 2012;30:350-359.

54. Timmins NE, Palfreyman E, Marturana F, Dietmair S, Luikenga S, Lopez G, Fung YL, et al. Clinical scale ex vivo manufacture of neutrophils from hematopoietic progenitor cells. Biotechnol Bioeng 2009;104:832-840.

55. King JA, Miller WM. Bioreactor development for stem cell expansion and controlled differentiation. Curr Opin Chem Biol 2007;11:394-398.

56. Yang S, Cai H, Jin H, Tan WS. Hematopoietic reconstitution of CD34+ cells grown in static and stirred culture systems in NOD/SCID mice. Biotechnol Lett 2008;30:61-65.

57. Watanabe K, Ueno M, Kamiya D, Nishiyama A, Matsumura M, Wataya T, Takahashi JB, et al. A ROCK inhibitor permits survival of dissociated human embryonic stem cells. Nat Biotechnol 2007;25:681-686.

58. Santos F, Andrade PZ, Abecasis MM, Gimble JM, Chase LG, Campbell AM, Boucher S, et al. Toward a clinical-grade expansion of mesenchymal stem cells from human sources: a microcarrier-based culture system under xeno-free conditions. Tissue Eng Part C Methods 2011;17:1201-1210.

59. Eibes G, dos Santos F, Andrade PZ, Boura JS, Abecasis MM, da Silva CL, Cabral JM. Maximizing the ex vivo expansion of human mesenchymal stem cells using a microcarrier-based stirred culture system. J Biotechnol 2010;146:194-197.

60. Frauenschuh S, Reichmann E, Ibold Y, Goetz PM, Sittinger M, Ringe J. A microcarrier-based cultivation system for expansion of primary mesenchymal stem cells. Biotechnol Prog 2007;23:187-193.

61. Schop D, van Dijkhuizen-Radersma R, Borgart E, Janssen FW, Rozemuller H, Prins HJ, de Bruijn JD. Expansion of human mesenchymal stromal cells on microcarriers: growth and metabolism. J Tissue Eng Regen Med 2010;4:131-140.

62. Fernandes-Platzgummer A, Diogo MM, Baptista RP, da Silva CL, Cabral JM. Scale-up of mouse embryonic stem cell expansion in stirred bioreactors. Biotechnol Prog 2011;27:1421-1432.

63. Krawetz R, Taiani JT, Liu S, Meng G, Li X, Kallos MS, Rancourt DE. Large-scale expansion of pluripotent human embryonic stem cells in stirred-suspension bioreactors. Tissue Eng Part C Methods 2010;16:573-582.

64. Olmer R, Lange A, Selzer S, Kasper C, Haverich A, Martin U, Zweigerdt R. Suspension culture of human pluripotent stem cells in controlled, stirred bioreactors. Tissue Eng Part C Methods 2012;18:772-784.

65. Kehoe DE, Jing D, Lock LT, Tzanakakis ES. Scalable stirred-suspension bioreactor culture of human pluripotent stem cells. Tissue Eng Part A 2010;16:405-421.

66. Amit M, Laevsky I, Miropolsky Y, Shariki K, Peri M, Itskovitz-Eldor J. Dynamic suspension culture for scalable expansion of undifferentiated human pluripotent stem cells. Nat Protoc 2011;6:572-579.

67. Kedong S, Xiubo F, Tianqing L, Macedo HM, LiLi J, Meiyun F, Fangxin S, et al. Simultaneous expansion and harvest of hematopoietic stem cells and mesenchymal stem cells derived from umbilical cord blood. J Mater Sci Mater Med 2010;21:3183-3193.

68. Yokomizo AY, Antoniazzi MM, Galdino PL, Azambuja N, Jr., Jorge SA, Pereira CA. Rabies virus production in high vero cell density cultures on macroporous microcarriers. Biotechnol Bioeng 2004;85:506-515.

69. Montagnon BJ, Fanget B, Vincent-Falquet JC. Industrial-scale production of inactivated poliovirus vaccine prepared by culture of Vero cells on microcarrier. Rev Infect Dis 1984;6 Suppl 2:S341-344.

70. Rourou S, van der Ark A, Majoul S, Trabelsi K, van der Velden T, Kallel H. A novel animal-component-free medium for rabies virus production in Vero cells grown on Cytodex 1 microcarriers in a stirred bioreactor. Appl Microbiol Biotechnol 2009;85:53-63.

71. van Hemert P, Kilburn DG, van Wezel AL. Homogeneous cultivation of animal cells for the production of virus and virus products. Biotechnol Bioeng 1969;11:875-885.

72. Kistner O, Barrett PN, Mundt W, Reiter M, Schober-Bendixen S, Dorner F. Development of a mammalian cell (Vero) derived candidate influenza virus vaccine. Vaccine 1998;16:960-968.

73. Diogo MM, da Silva CL, Cabral JM. Separation technologies for stem cell bioprocessing. Biotechnol Bioeng 2012;109:2699-2709.

74. Syed BA, Evans JB. Stem cell therapy market. Nat Rev Drug Discov 2013;12:185-186.

75. Mulukutla BC, Yongky A, Le T, Mashek DG, Hu WS. Regulation of Glucose Metabolism - A Perspective From Cell Bioprocessing. Trends Biotechnol 2016;34:638-651.

76. Hu WS, Aunins JG. Large-scale mammalian cell culture. Curr Opin Biotechnol 1997;8:148-153.

77. Uldry M, Ibberson M, Hosokawa M, Thorens B. GLUT2 is a high affinity glucosamine transporter. FEBS Lett 2002;524:199-203.

78. Raju R, Chau D, Cho DS, Park Y, Verfaillie CM, Hu WS. Cell Expansion During Directed Differentiation of Stem Cells Toward the Hepatic Lineage. Stem Cells Dev 2017;26:274-284.

79. DeLaForest A, Nagaoka M, Si-Tayeb K, Noto FK, Konopka G, Battle MA, Duncan SA. HNF4A is essential for specification of hepatic progenitors from human pluripotent stem cells. Development 2011;138:4143-4153.

80. Simpson IA, Dwyer D, Malide D, Moley KH, Travis A, Vannucci SJ. The facilitative glucose transporter GLUT3: 20 years of distinction. Am J Physiol Endocrinol Metab 2008;295:E242-253.

81. Chen L, Alam T, Johnson JH, Hughes S, Newgard CB, Unger RH. Regulation of beta-cell glucose transporter gene expression. Proc Natl Acad Sci U S A 1990;87:4088-4092.

82. Liu Z, Lei M, Jiang Y, Hao H, Chu L, Xu J, Luo M, et al. High glucose attenuates VEGF expression in rat multipotent adult progenitor cells in association with inhibition of JAK2/STAT3 signalling. J Cell Mol Med 2009;13:3427-3436.

83. Weil BR, Abarbanell AM, Herrmann JL, Wang Y, Meldrum DR. High glucose concentration in cell culture medium does not acutely affect human mesenchymal stem cell growth factor production or proliferation. Am J Physiol Regul Integr Comp Physiol 2009;296:R1735-1743.

84. Fenselau A, Kaiser D, Wallis K. Nucleoside requirements for the in vitro growth of bovine aortic endothelial cells. J Cell Physiol 1981;108:375-384.

85. Bester AC, Roniger M, Oren YS, Im MM, Sarni D, Chaoat M, Bensimon A, et al. Nucleotide deficiency promotes genomic instability in early stages of cancer development. Cell 2011;145:435-446.

86. Ruiz S, Lopez-Contreras AJ, Gabut M, Marion RM, Gutierrez-Martinez P, Bua S, Ramirez O, et al. Limiting replication stress during somatic cell reprogramming reduces genomic instability in induced pluripotent stem cells. Nat Commun 2015;6:8036.

87. Traut TW. Physiological concentrations of purines and pyrimidines. Mol Cell Biochem 1994;140:1-22.

88. Sancho-Bru P, Roelandt P, Narain N, Pauwelyn K, Notelaers T, Shimizu T, Ott M, et al. Directed differentiation of murine-induced pluripotent stem cells to functional hepatocyte-like cells. J Hepatol 2011;54:98-107.

89. Agarwal S, Holton KL, Lanza R. Efficient differentiation of functional hepatocytes from human embryonic stem cells. Stem Cells 2008;26:1117-1127.

90. Si-Tayeb K, Noto FK, Nagaoka M, Li J, Battle MA, Duris C, North PE, et al. Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. Hepatology 2010;51:297-305.

91. Sharma AD, Cantz T, Vogel A, Schambach A, Haridass D, Iken M, Bleidissel M, et al. Murine embryonic stem cell-derived hepatic progenitor cells engraft in recipient livers with limited capacity of liver tissue formation. Cell Transplant 2008;17:313-323.

92. Li F, Liu P, Liu C, Xiang D, Deng L, Li W, Wangensteen K, et al. Hepatoblast-like progenitor cells derived from embryonic stem cells can repopulate livers of mice. Gastroenterology 2010;139:2158-2169 e2158.

93. Zhao D, Chen S, Cai J, Guo Y, Song Z, Che J, Liu C, et al. Derivation and characterization of hepatic progenitor cells from human embryonic stem cells. PLoS One 2009;4:e6468.

94. Rossi JM, Dunn NR, Hogan BL, Zaret KS. Distinct mesodermal signals, including BMPs from the septum transversum mesenchyme, are required in combination for hepatogenesis from the endoderm. Genes Dev 2001;15:1998-2009.

95. Radzisheuskaya A, Chia Gle B, dos Santos RL, Theunissen TW, Castro LF, Nichols J, Silva JC. A defined Oct4 level governs cell state transitions of pluripotency entry and differentiation into all embryonic lineages. Nat Cell Biol 2013;15:579-590.

96. Yasunaga M, Tada S, Torikai-Nishikawa S, Nakano Y, Okada M, Jakt LM, Nishikawa S, et al. Induction and monitoring of definitive and visceral endoderm differentiation of mouse ES cells. Nat Biotechnol 2005;23:1542-1550.

97. Nava S, Westgren M, Jaksch M, Tibell A, Broome U, Ericzon BG, Sumitran-Holgersson S. Characterization of cells in the developing human liver. Differentiation 2005;73:249-260.

98. Bendall SC, Simonds EF, Qiu P, Amir el AD, Krutzik PO, Finck R, Bruggner RV, et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science 2011;332:687-696.

99. Qiu P, Simonds EF, Bendall SC, Gibbs KD, Jr., Bruggner RV, Linderman MD, Sachs K, et al. Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. Nat Biotechnol 2011;29:886-891.

100. Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol 2006;7:R100.

101. Chen J, Raymond K. Roles of rifampicin in drug-drug interactions: underlying molecular mechanisms involving the nuclear pregnane X receptor. Ann Clin Microbiol Antimicrob 2006;5:3.

102. Subramanian K, Owens DJ, Raju R, Firpo M, O'Brien TD, Verfaillie CM, Hu WS. Spheroid culture for enhanced differentiation of human embryonic stem cells to hepatocyte-like cells. Stem Cells Dev 2014;23:124-131.

103. Zorn AM, Wells JM. Vertebrate endoderm development and organ formation. Annu Rev Cell Dev Biol 2009;25:221-251.

104. Zaret KS. Regulatory phases of early liver development: paradigms of organogenesis. Nat Rev Genet 2002;3:499-512.

105. Lemaigre F, Zaret KS. Liver development update: new embryo models, cell lineage control, and morphogenesis. Curr Opin Genet Dev 2004;14:582-590.

106. Sen N, Mukherjee G, Sen A, Bendall SC, Sung P, Nolan GP, Arvin AM. Single-cell mass cytometry analysis of human tonsil T cell remodeling by varicella zoster virus. Cell Rep 2014;8:633-645.

107. Yao Y, Liu R, Shin MS, Trentalange M, Allore H, Nassar A, Kang I, et al. CyTOF supports efficient detection of immune cell subsets from small samples. J Immunol Methods 2014;415:1-5.

108. Chang Q, Ornatsky OI, Koch CJ, Chaudary N, Marie-Egyptienne DT, Hill RP, Tanner SD, et al. Single-cell measurement of the uptake, intratumoral distribution and cell cycle effects of cisplatin using mass cytometry. Int J Cancer 2015;136:1202-1209.

109. Ornatsky O, Bandura D, Baranov V, Nitz M, Winnik MA, Tanner S. Highly multiparametric analysis by mass cytometry. J Immunol Methods 2010;361:1-20.

110. Giesen C, Wang HA, Schapiro D, Zivanovic N, Jacobs A, Hattendorf B, Schuffler PJ, et al. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. Nat Methods 2014;11:417-422.

111. Lujan E, Zunder ER, Ng YH, Goronzy IN, Nolan GP, Wernig M. Early reprogramming regulators identified by prospective isolation and mass cytometry. Nature 2015;521:352-356.

112. Oertel M, Menthena A, Chen YQ, Teisner B, Jensen CH, Shafritz DA. Purification of fetal liver stem/progenitor cells containing all the repopulation potential for normal adult rat liver. Gastroenterology 2008;134:823-832.

113. Tanaka M, Okabe M, Suzuki K, Kamiya Y, Tsukahara Y, Saito S, Miyajima A. Mouse hepatoblasts at distinct developmental stages are characterized by expression of EpCAM and DLK1: drastic change of EpCAM expression during liver development. Mech Dev 2009;126:665-676.

114. Takebe T, Sekine K, Enomura M, Koike H, Kimura M, Ogaeri T, Zhang RR, et al. Vascularized and functional human liver from an iPSC-derived organ bud transplant. Nature 2013;499:481-484.

115. Soto-Gutierrez A, Navarro-Alvarez N, Zhao D, Rivas-Carrillo JD, Lebkowski J, Tanaka N, Fox IJ, et al. Differentiation of mouse embryonic stem cells to hepatocyte-like cells by co-culture with human liver nonparenchymal cell lines. Nat Protoc 2007;2:347-356.

116. Du C, Narayanan K, Leong MF, Wan AC. Induced pluripotent stem cell-derived hepatocytes and endothelial cells in multi-component hydrogel fibers for liver tissue engineering. Biomaterials 2014;35:6006-6014.

117. Siller R, Greenhough S, Naumovska E, Sullivan GJ. Small-molecule-driven hepatocyte differentiation of human pluripotent stem cells. Stem Cell Reports 2015;4:939-952.

118. Tasnim F, Phan D, Toh YC, Yu H. Cost-effective differentiation of hepatocyte-like cells from human pluripotent stem cells using small molecules. Biomaterials 2015;70:115-125.

119. Shan J, Schwartz RE, Ross NT, Logan DJ, Thomas D, Duncan SA, North TE, et al. Identification of small molecules for human hepatocyte expansion and iPS differentiation. Nat Chem Biol 2013;9:514-520.

120. Takayama K, Inamura M, Kawabata K, Sugawara M, Kikuchi K, Higuchi M, Nagamoto Y, et al. Generation of metabolically functioning hepatocytes from human pluripotent stem cells by FOXA2 and HNF1alpha transduction. J Hepatol 2012;57:628-636.

121. Takayama K, Inamura M, Kawabata K, Katayama K, Higuchi M, Tashiro K, Nonaka A, et al. Efficient Generation of Functional Hepatocytes From Human Embryonic Stem Cells and Induced Pluripotent Stem Cells by HNF4alpha Transduction. Mol Ther 2012;20:127-137.

122. Du Y, Wang J, Jia J, Song N, Xiang C, Xu J, Hou Z, et al. Human hepatocytes with drug metabolic function induced from fibroblasts by lineage reprogramming. Cell Stem Cell 2014;14:394-403.

123. Godoy P, Schmidt-Heck W, Natarajan K, Lucendo-Villarin B, Szkolnicka D, Asplund A, Bjorquist P, et al. Gene networks and transcription factor motifs defining the differentiation of stem cells into hepatocyte-like cells. J Hepatol 2015;63:934-942.

124. Baxter M, Withey S, Harrison S, Segeritz CP, Zhang F, Atkinson-Dell R, Rowe C, et al. Phenotypic and functional analyses show stem cell-derived hepatocyte-like cells better mimic fetal rather than adult hepatocytes. J Hepatol 2015;62:581-589.

125. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 2007;8:118-127.

126. Nygaard V, Rodland EA, Hovig E. Methods that remove batch effects while retaining group differences may lead to exaggerated confidence in downstream analyses. Biostatistics 2016;17:29-39.

127. Benito M, Parker J, Du Q, Wu J, Xiang D, Perou CM, Marron JS. Adjustment of systematic microarray data biases. Bioinformatics 2004;20:105-114.

128. Sims AH, Smethurst GJ, Hey Y, Okoniewski MJ, Pepper SD, Howell A, Miller CJ, et al. The removal of multiplicative, systematic bias allows integration of breast cancer gene expression datasets - improving meta-analysis and prediction of prognosis. BMC Med Genomics 2008;1:42.

129. Chen C, Grennan K, Badner J, Zhang D, Gershon E, Jin L, Liu C. Removing batch effects in analysis of expression microarray data: an evaluation of six batch adjustment methods. PLoS One 2011;6:e17238.

130. Luo J, Schumacher M, Scherer A, Sanoudou D, Megherbi D, Davison T, Shi T, et al. A comparison of batch effect removal methods for enhancement of prediction performance using MAQC-II microarray gene expression data. Pharmacogenomics J 2010;10:278-291.

131. Godoy P, Schmidt-Heck W, Natarajan K, Lucendo-Villarin B, Szkolnicka D, Asplund A, Bjorquist P, et al. Gene networks and transcription factor motifs defining the differentiation of stem cells into hepatocyte-like cells. J Hepatol 2015.

132. Jozefczuk J, Prigione A, Chavez L, Adjaye J. Comparative analysis of human embryonic stem cell and induced pluripotent stem cell-derived hepatocyte-like cells reveals current drawbacks and possible strategies for improved differentiation. Stem Cells Dev 2011;20:1259-1275.

133. Roelandt P, Obeid S, Paeshuyse J, Vanhove J, Van Lommel A, Nahmias Y, Nevens F, et al. Human pluripotent stem cell-derived hepatocytes support complete replication of hepatitis C virus. J Hepatol 2012;57:246-251.

134. Touboul T, Hannan NR, Corbineau S, Martinez A, Martinet C, Branchereau S, Mainot S, et al. Generation of functional hepatocytes from human embryonic stem cells under chemically defined conditions that recapitulate liver development. Hepatology 2010;51:1754-1765.

135. Li T, Huang J, Jiang Y, Zeng Y, He F, Zhang MQ, Han Z, et al. Multi-stage analysis of gene expression and transcription regulation in C57/B6 mouse liver development. Genomics 2009;93:235-242.

136. Alter O, Brown PO, Botstein D. Singular value decomposition for genome-wide expression data processing and modeling. Proc Natl Acad Sci U S A 2000;97:10101-10106.

137. Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. Proc Natl Acad Sci U S A 2004;101:4164-4169.

138. Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. BMC Bioinformatics 2010;11:367.

139. Kim H, Park H. Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. Bioinformatics 2007;23:1495-1502.

140. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 2009;37:1-13.

141. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44-57.

142. Nelms BL, Labosky PA: In: Transcriptional Control of Neural Crest Development. San Rafael (CA), 2010.

143. Strumpf D, Mao CA, Yamanaka Y, Ralston A, Chawengsaksophak K, Beck F, Rossant J. Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development 2005;132:2093-2102.

144. Gao N, White P, Kaestner KH. Establishment of intestinal identity and epithelial-mesenchymal signaling by Cdx2. Dev Cell 2009;16:588-599.

145. McFadden DG, Barbosa AC, Richardson JA, Schneider MD, Srivastava D, Olson EN. The Hand1 and Hand2 transcription factors regulate expansion of the embryonic cardiac ventricles in a gene dosage-dependent manner. Development 2005;132:189-201.

146. Duboc V, Logan MP. Pitx1 is necessary for normal initiation of hindlimb outgrowth through regulation of Tbx4 expression and shapes hindlimb morphologies via targeted growth control. Development 2011;138:5301-5309.

147. Bensoussan-Trigano V, Lallemand Y, Saint Cloment C, Robert B. Msx1 and Msx2 in limb mesenchyme modulate digit number and identity. Dev Dyn 2011;240:1190-1202.

148. Belo J, Krishnamurthy M, Oakie A, Wang R. The role of SOX9 transcription factor in pancreatic and duodenal development. Stem Cells Dev 2013;22:2935-2943.

149. Casas E, Kim J, Bendesky A, Ohno-Machado L, Wolfe CJ, Yang J. Snail2 is an essential mediator of Twist1-induced epithelial mesenchymal transition and metastasis. Cancer Res 2011;71:245-254.

150. Fenouille N, Tichet M, Dufies M, Pottier A, Mogha A, Soo JK, Rocchi S, et al. The epithelial-mesenchymal transition (EMT) regulatory factor SLUG (SNAI2) is a downstream target of SPARC and AKT in promoting melanoma cell invasion. PLoS One 2012;7:e40378.

151. Raum JC, Soleimanpour SA, Groff DN, Core N, Fasano L, Garratt AN, Dai C, et al. Tshz1 Regulates Pancreatic beta-Cell Maturation. Diabetes 2015;64:2905-2914.

152. Cain JT, Berosik MA, Snyder SD, Crawford NF, Nour SI, Schaubhut GJ, Darland DC. Shifts in the vascular endothelial growth factor isoforms result in transcriptome changes correlated with early neural stem cell proliferation and differentiation in mouse forebrain. Dev Neurobiol 2014;74:63-81.

153. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, Nair VS, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat Med 2015;21:938-945.

154. Ma H, Morey R, O'Neil RC, He Y, Daughtry B, Schultz MD, Hariharan M, et al. Abnormalities in human pluripotent cells due to reprogramming mechanisms. Nature 2014;511:177-183.

155. Miller JA, Ding SL, Sunkin SM, Smith KA, Ng L, Szafer A, Ebbert A, et al. Transcriptional landscape of the prenatal human brain. Nature 2014;508:199-206.

156. Muller C, Schillert A, Rothemeier C, Tregouet DA, Proust C, Binder H, Pfeiffer N, et al. Removing Batch Effects from Longitudinal Gene Expression - Quantile Normalization Plus ComBat as Best Approach for Microarray Transcriptome Data. PLoS One 2016;11:e0156594.

157. Rempel E, Hoelting L, Waldmann T, Balmer NV, Schildknecht S, Grinberg M, Das Gaspar JA, et al. A transcriptome-based classifier to identify developmental toxicants by stem cell testing: design, validation and optimization for histone deacetylase inhibitors. Arch Toxicol 2015;89:1599-1618.

158. Downs KM, Davies T. Staging of gastrulating mouse embryos by morphological landmarks in the dissecting microscope. Development 1993;118:1255-1266.

159. Clancy B, Darlington RB, Finlay BL. Translating developmental time across mammalian species. Neuroscience 2001;105:7-17.

160. Levin M, Hashimshony T, Wagner F, Yanai I. Developmental milestones punctuate gene expression in the Caenorhabditis embryo. Dev Cell 2012;22:1101-1108.

161. Chitwood DH, Headland LR, Kumar R, Peng J, Maloof JN, Sinha NR. The developmental trajectory of leaflet morphology in wild tomato species. Plant Physiol 2012;158:1230-1240.

162. Anavy L, Levin M, Khair S, Nakanishi N, Fernandez-Valverde SL, Degnan BM, Yanai I. BLIND ordering of large-scale transcriptomic developmental timecourses. Development 2014;141:1161-1166.

163. Forristal CE, Wright KL, Hanley NA, Oreffo RO, Houghton FD. Hypoxia inducible factors regulate pluripotency and proliferation in human embryonic stem cells cultured at reduced oxygen tensions. Reproduction 2010;139:85-97.

164. Ezashi T, Das P, Roberts RM. Low O2 tensions and the prevention of differentiation of hES cells. Proc Natl Acad Sci U S A 2005;102:4783-4788.

165. Pan C, Kumar C, Bohl S, Klingmueller U, Mann M. Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific functions. Mol Cell Proteomics 2009;8:443-450.

166. Chan C, Berthiaume F, Lee K, Yarmush ML. Metabolic flux analysis of cultured hepatocytes exposed to plasma. Biotechnol Bioeng 2003;81:33-49.

167. Chan C, Berthiaume F, Washizu J, Toner M, Yarmush ML. Metabolic pre-conditioning of cultured cells in physiological levels of insulin: generating resistance to the lipid-accumulating effects of plasma in hepatocytes. Biotechnol Bioeng 2002;78:753-760.

168. Varum S, Rodrigues AS, Moura MB, Momcilovic O, Easley CAt, Ramalho-Santos J, Van Houten B, et al. Energy metabolism in human pluripotent stem cells and their differentiated counterparts. PLoS One 2011;6:e20914.

169. Folmes CD, Dzeja PP, Nelson TJ, Terzic A. Metabolic plasticity in stem cell homeostasis and differentiation. Cell Stem Cell 2012;11:596-606.

170. Roelandt P, Pauwelyn KA, Sancho-Bru P, Subramanian K, Bose B, Ordovas L, Vanuytsel K, et al. Human embryonic and rat adult stem cells with primitive endoderm-like phenotype can be fated to definitive endoderm, and finally hepatocyte-like cells. PLoS One 2010;5:e12101.

171. Higuera GA, Schop D, Spitters TW, van Dijkhuizen-Radersma R, Bracke M, de Bruijn JD, Martens D, et al. Patterns of amino acid metabolism by proliferating human mesenchymal stem cells. Tissue Eng Part A 2012;18:654-664.

172. Shatford RA, Nyberg SL, Meier SJ, White JG, Payne WD, Hu WS, Cerra FB. Hepatocyte function in a hollow fiber bioreactor: a potential bioartificial liver. J Surg Res 1992;53:549-557.

173. Nyberg SL, Remmel RP, Mann HJ, Peshwa MV, Hu WS, Cerra FB. Primary hepatocytes outperform Hep G2 cells as the source of biotransformation functions in a bioartificial liver. Ann Surg 1994;220:59-67.

174. Sielaff TD, Nyberg SL, Rollins MD, Hu MY, Amiot B, Lee A, Wu FJ, et al. Characterization of the three-compartment gel-entrapment porcine hepatocyte bioartificial liver. Cell Biol Toxicol 1997;13:357-364.

175. Smith MD, Smirthwaite AD, Cairns DE, Cousins RB, Gaylor JD. Techniques for measurement of oxygen consumption rates of hepatocytes during attachment and post-attachment. Int J Artif Organs 1996;19:36-44.

176. Mandal S, Lindgren AG, Srivastava AS, Clark AT, Banerjee U. Mitochondrial function controls proliferation and early differentiation potential of embryonic stem cells. Stem Cells 2011;29:486-495.

177. Chung S, Dzeja PP, Faustino RS, Perez-Terzic C, Behfar A, Terzic A. Mitochondrial oxidative metabolism is required for the cardiac differentiation of stem cells. Nat Clin Pract Cardiovasc Med 2007;4 Suppl 1:S60-67.

178. Porter GA, Jr., Hom J, Hoffman D, Quintanilla R, de Mesy Bentley K, Sheu SS. Bioenergetics, mitochondria, and cardiac myocyte differentiation. Prog Pediatr Cardiol 2011;31:75-81.

179. Rehman J. Empowering self-renewal and differentiation: the role of mitochondria in stem cells. J Mol Med (Berl) 2010;88:981-986.

180. Armstrong L, Tilgner K, Saretzki G, Atkinson SP, Stojkovic M, Moreno R, Przyborski S, et al. Human induced pluripotent stem cell lines show stress defense mechanisms and mitochondrial regulation similar to those of human embryonic stem cells. Stem Cells 2010;28:661-673.

181. Rathmell JC, Fox CJ, Plas DR, Hammerman PS, Cinalli RM, Thompson CB. Akt-directed glucose metabolism can prevent Bax conformation change and promote growth factor-independent survival. Mol Cell Biol 2003;23:7315-7328.

182. Freemerman AJ, Johnson AR, Sacks GN, Milner JJ, Kirk EL, Troester MA, Macintyre AN, et al. Metabolic reprogramming of macrophages: glucose transporter 1 (GLUT1)-mediated glucose metabolism drives a proinflammatory phenotype. J Biol Chem 2014;289:7884-7896.

183. Gray LR, Rauckhorst AJ, Taylor EB. A Method for Multiplexed Measurement of Mitochondrial Pyruvate Carrier Activity. J Biol Chem 2016;291:7409-7417.

184. McCommis KS, Chen Z, Fu X, McDonald WG, Colca JR, Kletzien RF, Burgess SC, et al. Loss of Mitochondrial Pyruvate Carrier 2 in the Liver Leads to Defects in Gluconeogenesis and Compensation via Pyruvate-Alanine Cycling. Cell Metab 2015;22:682-694.

Chapter 9. Appendix

Table 9-1. Metagenes Identified from Human *In-vitro* NMF Analysis

Gene ID	Metagene
ENSG00000154277	1
ENSG00000138829	1
ENSG00000158270	1
ENSG00000043355	1
ENSG00000198467	1
ENSG00000131914	1
ENSG00000213120	1
ENSG00000187772	1
ENSG00000130294	1
ENSG00000104332	1
ENSG00000163629	1
ENSG00000163508	1
ENSG00000175928	1
ENSG00000006118	1
ENSG00000137285	1
ENSG00000121966	1
ENSG00000156925	1
ENSG00000106278	1
ENSG00000170373	1
ENSG00000090530	1
ENSG00000171208	1
ENSG00000151640	1
ENSG00000116729	1
ENSG00000095596	1
ENSG00000184867	1
ENSG00000268151	1
ENSG00000174469	1
ENSG00000105270	1
ENSG00000158246	1
ENSG00000123560	1
ENSG00000136943	1
ENSG00000143768	1
ENSG00000164434	1
ENSG00000168280	1
ENSG00000262907	1
ENSG00000067445	1
ENSG00000082497	1

ENSG00000088756	1
ENSG00000177519	1
ENSG00000184261	1
ENSG00000026025	1
ENSG00000111716	1
ENSG00000141971	1
ENSG00000166503	1
ENSG00000178531	1
ENSG00000213190	1
ENSG00000176887	1
ENSG00000092445	1
ENSG00000243709	1
ENSG00000167157	1
ENSG00000121570	1
ENSG00000166086	1
ENSG00000128045	1
ENSG00000125285	1
ENSG00000169169	1
ENSG00000143320	1
ENSG00000050165	1
ENSG00000075213	1
ENSG00000114854	1
ENSG00000198211	1
ENSG00000258947	1
ENSG00000106631	1
ENSG00000131711	1
ENSG00000104728	1
ENSG00000168268	1
ENSG00000138650	1
ENSG00000076716	1
ENSG00000082397	1
ENSG00000179046	1
ENSG00000132688	1
ENSG00000149591	1
ENSG00000184304	1
ENSG00000196368	1
ENSG00000269203	1
ENSG00000041982	1
ENSG00000148143	1
ENSG00000109089	1
ENSG00000109705	1
ENSG00000147869	1
ENSG00000138180	1

ENSG00000046653	1
ENSG00000114948	1
ENSG00000105655	1
ENSG00000165349	1
ENSG00000162551	1
ENSG00000133636	1
ENSG00000181449	1
ENSG00000165588	1
ENSG00000134775	1
ENSG00000152377	1
ENSG00000196526	1
ENSG00000103449	1
ENSG00000164659	1
ENSG00000144339	1
ENSG00000113758	1
ENSG00000144834	1
ENSG00000165495	1
ENSG00000149201	1
ENSG00000240694	1
ENSG00000074047	1
ENSG00000136231	1
ENSG00000139946	1
ENSG00000124766	1
ENSG00000272398	1
ENSG00000126878	1
ENSG00000088882	1
ENSG00000019144	1
ENSG00000267867	1
ENSG00000184613	1
ENSG00000134531	1
ENSG00000172348	1
ENSG00000061337	1
ENSG00000026559	1
ENSG00000204335	1
ENSG00000183145	1
ENSG00000143632	1
ENSG00000104722	1
ENSG00000240563	1
ENSG00000128683	1
ENSG00000159167	1
ENSG00000138795	1
ENSG00000109255	1
ENSG00000166426	1

ENSG00000134323	1
ENSG00000198795	1
ENSG00000137821	1
ENSG00000204711	1
ENSG00000033170	1
ENSG00000196376	1
ENSG00000150551	1
ENSG00000161249	1
ENSG00000157827	1
ENSG00000101144	1
ENSG00000153707	1
ENSG00000144810	1
ENSG00000128596	1
ENSG00000122861	1
ENSG00000169992	1
ENSG00000179431	1
ENSG00000154188	1
ENSG00000160326	1
ENSG00000261768	1
ENSG00000128567	1
ENSG00000167657	1
ENSG00000145321	2
ENSG00000136872	2
ENSG00000117601	2
ENSG00000187758	2
ENSG00000118520	2
ENSG00000106927	2
ENSG00000122194	2
ENSG00000091583	2
ENSG00000099937	2
ENSG00000229314	2
ENSG00000113889	2
ENSG00000101981	2
ENSG00000123561	2
ENSG00000110245	2
ENSG00000135744	2
ENSG00000138207	2
ENSG00000171557	2
ENSG00000145192	2
ENSG00000091513	2
ENSG00000163631	2
ENSG00000197249	2
ENSG00000171560	2

ENSG00000086696	2
ENSG00000080618	2
ENSG00000156096	2
ENSG00000180432	2
ENSG00000151655	2
ENSG00000261701	2
ENSG00000198099	2
ENSG00000163581	2
ENSG00000110169	2
ENSG00000079557	2
ENSG00000090512	2
ENSG00000036473	2
ENSG00000088926	2
ENSG00000105398	2
ENSG00000205364	2
ENSG00000211452	2
ENSG00000172482	2
ENSG00000151224	2
ENSG00000114771	2
ENSG00000104760	2
ENSG00000169903	2
ENSG00000066813	2
ENSG00000157131	2
ENSG00000145692	2
ENSG00000131482	2
ENSG00000171759	2
ENSG00000021852	2
ENSG00000106565	2
ENSG00000249948	2
ENSG00000080910	2
ENSG00000257017	2
ENSG00000072080	2
ENSG00000168509	2
ENSG00000265970	2
ENSG00000164344	2
ENSG00000166035	2
ENSG00000136881	2
ENSG00000106258	2
ENSG00000130173	2
ENSG00000160862	2
ENSG00000134240	2
ENSG00000002933	2
ENSG00000132703	2

ENSG00000118514	2
ENSG00000105697	2
ENSG00000165471	2
ENSG00000145826	2
ENSG00000100652	2
ENSG00000161944	2
ENSG00000162267	2
ENSG00000055957	2
ENSG00000039537	2
ENSG00000160870	2
ENSG00000142484	2
ENSG00000148584	2
ENSG00000147647	2
ENSG00000165092	2
ENSG00000178772	2
ENSG00000158104	2
ENSG00000112337	2
ENSG00000101323	2
ENSG00000113924	2
ENSG00000171766	2
ENSG00000105852	2
ENSG00000164406	2
ENSG00000055955	2
ENSG00000165140	2
ENSG00000129988	2
ENSG00000125144	2
ENSG00000012504	2
ENSG00000198734	2
ENSG00000244734	2
ENSG00000084674	2
ENSG00000167711	2
ENSG00000164089	2
ENSG00000099769	2
ENSG00000248144	2
ENSG00000084110	2
ENSG00000153086	2
ENSG00000140093	2
ENSG00000007933	2
ENSG00000138109	2
ENSG00000111700	2
ENSG00000198650	2
ENSG00000180745	2
ENSG00000188488	2

ENSG00000175336	2
ENSG00000113492	2
ENSG00000132693	2
ENSG00000105707	2
ENSG00000125551	2
ENSG00000183281	2
ENSG00000134962	2
ENSG00000174567	2
ENSG00000177575	2
ENSG00000170458	2
ENSG00000141485	2
ENSG00000179761	2
ENSG00000133488	2
ENSG00000009724	2
ENSG00000173369	2
ENSG00000151790	2
ENSG00000262635	2
ENSG00000172955	2
ENSG00000170099	2
ENSG00000132840	2
ENSG00000139344	2
ENSG00000166278	2
ENSG00000204364	2
ENSG00000206372	2
ENSG00000226560	2
ENSG00000231543	2
ENSG00000235017	2
ENSG00000235696	2
ENSG00000166816	2
ENSG00000165828	2
ENSG00000204359	2
ENSG00000239754	2
ENSG00000241253	2
ENSG00000241534	2
ENSG00000242335	2
ENSG00000243570	2
ENSG00000243649	2
ENSG00000125730	2
ENSG00000047457	2
ENSG00000139194	2
ENSG00000204287	2
ENSG00000226260	2
ENSG00000228987	2

ENSG00000230726	2
ENSG00000234794	2
ENSG00000227993	2
ENSG00000206308	2
ENSG00000109758	2
ENSG00000134365	2
ENSG00000108242	2
ENSG00000114200	2
ENSG00000115718	2
ENSG00000138823	2
ENSG00000137491	2
ENSG00000116882	2
ENSG00000172497	2
ENSG00000129151	2
ENSG00000135423	2
ENSG00000123838	2
ENSG00000176919	2
ENSG00000163959	2
ENSG00000110243	2
ENSG00000069535	2
ENSG00000166840	2
ENSG00000000971	2
ENSG00000143278	2
ENSG00000148702	2
ENSG00000188582	2
ENSG00000143546	2
ENSG00000130988	2
ENSG00000270349	2
ENSG00000171236	2
ENSG00000162365	2
ENSG00000187048	2
ENSG00000146678	2
ENSG00000159189	2
ENSG00000163825	2
ENSG00000131187	2
ENSG00000154262	2
ENSG00000021826	2
ENSG00000140284	2
ENSG00000157103	2
ENSG00000169562	2
ENSG00000123843	2
ENSG00000138075	2
ENSG00000111249	2

ENSG00000175899	2
ENSG00000198610	2
ENSG00000266359	2
ENSG00000077420	2
ENSG00000011600	2
ENSG00000023839	2
ENSG00000144908	2
ENSG00000100665	2
ENSG00000144891	2
ENSG00000134716	2
ENSG00000137204	2
ENSG00000109511	2
ENSG00000196136	2
ENSG00000273259	2
ENSG00000138115	2
ENSG00000198670	2
ENSG00000135447	2
ENSG00000126231	2
ENSG00000166927	2
ENSG00000103569	2
ENSG00000100024	2
ENSG00000122862	2
ENSG00000117707	2
ENSG00000139547	2
ENSG00000136305	2
ENSG00000169856	2
ENSG00000158874	2
ENSG00000245848	2
ENSG00000144837	2
ENSG00000019169	2
ENSG00000131910	2
ENSG00000112299	2
ENSG00000149124	2
ENSG00000005471	2
ENSG00000111181	2
ENSG00000198300	2
ENSG00000160282	2
ENSG00000142748	2
ENSG00000120915	2
ENSG00000108924	2
ENSG00000198692	2
ENSG00000113600	2
ENSG00000196620	2

ENSG00000258085	2
ENSG00000214274	2
ENSG00000186204	2
ENSG00000198417	2
ENSG00000113905	2
ENSG00000073734	2
ENSG00000263298	2
ENSG00000124568	2
ENSG00000138030	2
ENSG00000116791	2
ENSG00000060566	2
ENSG00000132855	2
ENSG00000204099	2
ENSG00000187045	2
ENSG00000124713	2
ENSG00000131781	2
ENSG00000266748	2
ENSG00000148346	2
ENSG00000174990	2
ENSG00000205403	2
ENSG00000143819	2
ENSG00000165682	2
ENSG00000025423	2
ENSG00000009950	2
ENSG00000262077	2
ENSG00000170439	2
ENSG00000215644	2
ENSG00000169174	2
ENSG00000171903	2
ENSG00000182902	2
ENSG00000163687	2
ENSG00000110077	2
ENSG00000120054	2
ENSG00000124253	2
ENSG00000083807	2
ENSG00000250799	2
ENSG00000135218	2
ENSG00000141505	2
ENSG00000266964	2
ENSG00000144035	2
ENSG00000132437	2
ENSG00000182326	2
ENSG00000269882	2

ENSG00000087237	2
ENSG00000163347	2
ENSG00000132541	2
ENSG00000126218	2
ENSG00000186115	2
ENSG00000196660	2
ENSG00000175003	2
ENSG00000160868	2
ENSG00000130208	2
ENSG00000138308	2
ENSG00000108515	2
ENSG00000196139	2
ENSG00000265685	2
ENSG00000187193	2
ENSG00000159423	2
ENSG00000117009	2
ENSG00000213424	2
ENSG00000264058	2
ENSG00000113790	2
ENSG00000169877	2
ENSG00000151365	2
ENSG00000111796	2
ENSG00000197444	2
ENSG00000196600	2
ENSG00000121858	2
ENSG00000109819	2
ENSG00000167798	2
ENSG00000140107	2
ENSG00000152804	2
ENSG00000168384	2
ENSG00000206291	2
ENSG00000224103	2
ENSG00000228163	2
ENSG00000229685	2
ENSG00000231389	2
ENSG00000235844	2
ENSG00000236177	2
ENSG00000161573	2
ENSG00000183549	2
ENSG00000115919	2
ENSG00000134538	2
ENSG00000141293	2
ENSG00000171840	2

ENSG00000149742	2
ENSG00000117594	2
ENSG00000154274	2
ENSG00000100197	2
ENSG00000272000	2
ENSG00000272532	2
ENSG00000125246	2
ENSG00000106804	2
ENSG00000160255	2
ENSG00000133574	2
ENSG00000134709	2
ENSG00000157873	2
ENSG00000136011	2
ENSG00000171747	2
ENSG00000021488	2
ENSG00000204444	2
ENSG00000206409	2
ENSG00000224290	2
ENSG00000226215	2
ENSG00000227567	2
ENSG00000231974	2
ENSG00000235754	2
ENSG00000062282	2
ENSG00000121691	2
ENSG00000146233	2
ENSG00000241935	2
ENSG00000106538	2
ENSG00000187017	2
ENSG00000084734	2
ENSG00000143845	2
ENSG00000139144	2
ENSG00000189221	2
ENSG00000151726	2
ENSG00000106327	2
ENSG00000118513	2
ENSG00000244067	2
ENSG00000145850	2
ENSG00000162817	2
ENSG00000258818	2
ENSG00000259171	2
ENSG00000170608	2
ENSG00000110887	2
ENSG00000127951	2

ENSG00000155659	2
ENSG00000129596	2
ENSG00000204291	3
ENSG00000198759	3
ENSG00000139329	3
ENSG00000113083	3
ENSG00000087303	3
ENSG00000113196	3
ENSG00000118271	3
ENSG00000164266	3
ENSG00000081051	3
ENSG00000174498	3
ENSG00000171564	3
ENSG00000173391	3
ENSG00000079112	3
ENSG00000134258	3
ENSG00000164294	3
ENSG00000170290	3
ENSG00000040731	3
ENSG00000140937	3
ENSG00000130635	3
ENSG00000164764	3
ENSG00000125872	3
ENSG00000107165	3
ENSG00000115380	3
ENSG00000128591	3
ENSG00000185559	3
ENSG00000164825	3
ENSG00000162998	3
ENSG00000261258	3
ENSG00000141756	3
ENSG00000069011	3
ENSG00000118495	3
ENSG00000169116	3
ENSG00000157005	3
ENSG00000168743	3
ENSG00000114251	3
ENSG00000110811	3
ENSG00000268828	3
ENSG00000019549	3
ENSG00000167874	3
ENSG00000139055	3
ENSG00000138119	3

ENSG00000112818	3
ENSG00000164484	3
ENSG00000113657	3
ENSG00000183160	3
ENSG00000164125	3
ENSG00000104415	3
ENSG00000270132	3
ENSG00000135919	3
ENSG00000145147	3
ENSG00000182752	3
ENSG00000126016	3
ENSG00000084636	3
ENSG00000138615	3
ENSG00000148677	3
ENSG00000203857	3
ENSG00000164692	3
ENSG00000104368	3
ENSG00000163520	3
ENSG00000168952	3
ENSG00000011465	3
ENSG00000135046	3
ENSG00000163359	3
ENSG00000129757	3
ENSG00000007062	3
ENSG00000103241	3
ENSG00000143125	3
ENSG00000159212	3
ENSG00000171004	3
ENSG00000060718	3
ENSG00000132031	3
ENSG00000171345	3
ENSG00000163430	3
ENSG00000180340	3
ENSG00000110244	3
ENSG00000170955	3
ENSG00000168542	3
ENSG00000105825	3
ENSG00000149090	3
ENSG00000099960	3
ENSG00000166147	3
ENSG00000173068	3
ENSG00000135480	3
ENSG00000145824	3

ENSG00000174807	3
ENSG00000218336	3
ENSG00000073756	3
ENSG00000154856	3
ENSG00000077942	3
ENSG00000169129	3
ENSG00000162493	3
ENSG00000108821	3
ENSG00000113739	3
ENSG00000086991	3
ENSG00000077616	3
ENSG00000111319	3
ENSG00000136193	3
ENSG00000198796	3
ENSG00000170425	3
ENSG00000164116	3
ENSG00000037965	3
ENSG00000197614	3
ENSG00000125384	3
ENSG00000152661	3
ENSG00000120937	3
ENSG00000188176	3
ENSG00000167617	3
ENSG00000261931	3
ENSG00000263030	3
ENSG00000270365	3
ENSG00000167434	3
ENSG00000133110	3
ENSG00000137573	3
ENSG00000141526	3
ENSG00000130176	3
ENSG00000067057	3
ENSG00000102359	3
ENSG00000150556	3
ENSG00000110660	3
ENSG00000163132	3
ENSG00000144645	3
ENSG00000168386	3
ENSG00000163624	3
ENSG00000164093	3
ENSG00000171033	3
ENSG00000101198	3
ENSG00000103196	3

ENSG00000100739	3
ENSG00000260225	3
ENSG00000112852	3
ENSG00000165556	3
ENSG00000071282	3
ENSG00000130508	3
ENSG00000141696	3
ENSG00000261163	3
ENSG00000176788	3
ENSG00000167755	3
ENSG00000120068	3
ENSG00000260027	3
ENSG00000171729	3
ENSG00000143867	3
ENSG00000162576	3
ENSG00000148344	3
ENSG00000173376	3
ENSG00000062038	3
ENSG00000105894	3
ENSG00000166250	3
ENSG00000171951	3
ENSG00000107485	3
ENSG00000122176	3
ENSG00000155465	3
ENSG00000133519	3
ENSG00000168824	3
ENSG00000070019	3
ENSG00000109099	3
ENSG00000172638	3
ENSG00000113140	3
ENSG00000013588	3
ENSG00000079215	3
ENSG00000138193	3
ENSG00000145681	3
ENSG00000104435	3
ENSG00000134871	3
ENSG00000172458	3
ENSG00000117525	3
ENSG00000136859	3
ENSG00000159251	3
ENSG00000118971	3
ENSG00000006747	3
ENSG00000117152	3

ENSG00000020181	3
ENSG00000140945	3
ENSG00000038427	3
ENSG00000197406	3
ENSG00000106031	3
ENSG00000153822	3
ENSG00000177875	3
ENSG00000154127	3
ENSG00000166482	3
ENSG00000150630	3
ENSG00000143036	3
ENSG00000079931	3
ENSG00000089472	3
ENSG00000114270	3
ENSG00000099256	3
ENSG00000166173	3
ENSG00000129038	3
ENSG00000139174	3
ENSG00000164932	3
ENSG00000111799	3
ENSG00000180318	3
ENSG00000150687	3
ENSG00000116774	3
ENSG00000145287	3
ENSG00000163947	3
ENSG00000146197	3
ENSG00000146250	3
ENSG00000124107	3
ENSG00000078098	3
ENSG00000183036	3
ENSG00000100079	3
ENSG00000101680	3
ENSG00000147655	3
ENSG00000204381	3
ENSG00000164251	3
ENSG00000120820	3
ENSG00000184515	3
ENSG00000269503	3
ENSG00000132854	3
ENSG00000166165	3
ENSG00000137269	3
ENSG00000132698	3
ENSG00000147883	3

ENSG00000163171	3
ENSG00000134369	3
ENSG00000130224	3
ENSG00000268417	3
ENSG00000079689	3
ENSG00000198108	3
ENSG00000179178	3
ENSG00000187123	3
ENSG00000123700	3
ENSG00000050555	3
ENSG00000197635	3
ENSG00000154734	3
ENSG00000137507	3
ENSG00000010932	3
ENSG00000064692	3
ENSG00000129116	3
ENSG00000100196	3
ENSG00000165794	3
ENSG00000169583	3
ENSG00000121297	3
ENSG00000133661	3
ENSG00000104723	3
ENSG00000168461	3
ENSG00000164106	3
ENSG00000184347	3
ENSG00000069702	3
ENSG00000091986	3
ENSG00000112964	3
ENSG00000169594	3
ENSG00000091136	3
ENSG00000101955	3
ENSG00000145555	3
ENSG00000167642	3
ENSG00000126947	3
ENSG00000269801	3
ENSG00000150938	3
ENSG00000261899	3
ENSG00000146374	3
ENSG00000164946	3
ENSG00000132915	3
ENSG00000138772	3
ENSG00000103064	3
ENSG00000198648	3

ENSG00000104044	3
ENSG00000109846	3
ENSG00000263007	3
ENSG00000125266	3
ENSG00000163637	3
ENSG00000118407	3
ENSG00000198832	3
ENSG00000046604	3
ENSG00000135549	3
ENSG00000100842	3
ENSG00000079257	3
ENSG00000167528	3
ENSG00000242779	3
ENSG00000152049	3
ENSG00000144619	3
ENSG00000166863	3
ENSG00000140682	3
ENSG00000174099	3
ENSG00000146674	3
ENSG00000164107	3
ENSG00000102034	3
ENSG00000164683	3
ENSG00000119888	3
ENSG00000163638	3
ENSG00000184005	3
ENSG00000165868	3
ENSG00000263052	3
ENSG00000122691	3
ENSG00000103175	3
ENSG00000206538	3
ENSG00000005884	3
ENSG00000099994	3
ENSG00000120149	3
ENSG00000015413	3
ENSG00000077943	3
ENSG00000189334	3
ENSG00000135736	3
ENSG00000163191	3
ENSG00000151612	3
ENSG00000125398	3
ENSG00000115604	3
ENSG00000139278	3
ENSG00000163288	3

ENSG00000166963	3
ENSG00000095752	3
ENSG00000101825	3
ENSG00000183876	3
ENSG00000142156	3
ENSG00000135052	3
ENSG00000116991	3
ENSG00000125850	3
ENSG00000163993	3
ENSG00000115602	3
ENSG00000161896	3
ENSG00000112655	3
ENSG00000184697	3
ENSG00000163017	3
ENSG00000115008	3
ENSG00000102243	3
ENSG00000197702	3
ENSG00000108244	3
ENSG00000263309	3
ENSG00000182534	3
ENSG00000119630	3
ENSG00000143839	3
ENSG00000106348	3
ENSG00000114113	3
ENSG00000166923	3
ENSG00000131459	3
ENSG00000119900	3
ENSG00000151693	3
ENSG00000157502	3
ENSG00000137809	3
ENSG00000188153	3
ENSG00000182718	3
ENSG00000109472	3
ENSG00000065911	3
ENSG00000168615	3
ENSG00000142552	3
ENSG00000132205	3
ENSG00000111339	3
ENSG00000141744	3
ENSG00000084207	3
ENSG00000158258	3
ENSG00000124102	3
ENSG00000140465	3

ENSG00000107317	3
ENSG00000172380	3
ENSG00000103888	3
ENSG00000125787	3
ENSG00000152137	3
ENSG00000104213	3
ENSG00000186377	3
ENSG00000167695	3
ENSG00000197380	3
ENSG00000078401	3
ENSG00000070404	3
ENSG00000171812	3
ENSG00000167644	3
ENSG00000197712	3
ENSG00000164124	3
ENSG00000198768	3
ENSG00000166394	3
ENSG00000153558	3
ENSG00000147003	3
ENSG00000144642	3
ENSG00000177984	3
ENSG00000081479	3
ENSG00000165379	3
ENSG00000052344	3
ENSG00000103742	3
ENSG00000198121	3
ENSG00000239474	3
ENSG00000240583	3
ENSG00000166145	3
ENSG00000175315	3
ENSG00000183722	3
ENSG00000142173	3
ENSG00000145730	3
ENSG00000165124	3
ENSG00000134668	3
ENSG00000133135	3
ENSG00000243978	3
ENSG00000182636	3
ENSG00000148942	3
ENSG00000129514	3
ENSG00000118503	3
ENSG00000118849	3
ENSG00000112175	3

ENSG00000146072	3
ENSG00000166900	3
ENSG00000105928	3
ENSG00000122585	3
ENSG00000174307	3
ENSG00000184838	3
ENSG00000173715	3
ENSG00000111913	3
ENSG00000118137	3
ENSG00000251493	3
ENSG00000204262	3
ENSG00000127472	3
ENSG00000142583	3
ENSG00000164318	3
ENSG00000160207	3
ENSG00000136026	3
ENSG00000132386	3
ENSG00000165105	3
ENSG00000050438	3
ENSG00000173559	3
ENSG00000170190	3
ENSG00000120708	3
ENSG00000181264	3
ENSG00000123096	3
ENSG00000152268	3
ENSG00000262655	3
ENSG00000183853	3
ENSG00000116017	3
ENSG00000182985	3
ENSG00000118194	3
ENSG00000001626	3
ENSG00000185565	3
ENSG00000135750	3
ENSG00000007384	3
ENSG00000049130	3
ENSG00000164120	3
ENSG00000147257	3
ENSG00000173546	3
ENSG00000108797	3
ENSG00000086548	3
ENSG00000272939	3
ENSG00000166920	3
ENSG00000130812	3

ENSG00000174482	3
ENSG00000139926	3
ENSG00000108947	3
ENSG00000179921	3
ENSG00000187498	3
ENSG00000110195	3
ENSG00000163362	3
ENSG00000156711	3
ENSG00000140545	3
ENSG00000183844	3
ENSG00000167535	3
ENSG00000181541	3
ENSG00000112562	3
ENSG00000156510	3
ENSG00000105989	3
ENSG00000142227	3
ENSG00000211448	3
ENSG00000129270	3
ENSG00000271447	3
ENSG00000185222	3
ENSG00000180447	3
ENSG00000116157	3
ENSG00000121440	3
ENSG00000168032	3
ENSG00000106366	3
ENSG00000176907	3
ENSG00000121361	3

Table 9-2. Metagenes Identified from Mouse In-vivo NMF Analysis.

Gene ID	Metagene
ENSMUSG00000026368	1
ENSMUSG00000031645	1
ENSMUSG00000038656	1
ENSMUSG00000054630	1
ENSMUSG00000029656	1
ENSMUSG00000029445	1
ENSMUSG00000070704	1
ENSMUSG00000040134	1
ENSMUSG00000022445	1
ENSMUSG00000062410	1
ENSMUSG00000037798	1
ENSMUSG00000027870	1
ENSMUSG00000030382	1
ENSMUSG00000026365	1
ENSMUSG00000028186	1
ENSMUSG00000069805	1
ENSMUSG00000031640	1
ENSMUSG00000027048	1
ENSMUSG00000060407	1
ENSMUSG00000025991	1
ENSMUSG00000022871	1
ENSMUSG00000074264	1
ENSMUSG00000074336	1
ENSMUSG00000026542	1
ENSMUSG00000030236	1
ENSMUSG00000022512	1
ENSMUSG00000057037	1
ENSMUSG00000030111	1
ENSMUSG00000002588	1
ENSMUSG00000056035	1
ENSMUSG00000021210	1
ENSMUSG00000056973	1
ENSMUSG00000031271	1
ENSMUSG00000026822	1
ENSMUSG00000038641	1
ENSMUSG00000037942	1
ENSMUSG00000033715	1
ENSMUSG00000027261	1
ENSMUSG00000024863	1

ENSMUSG00000031722	1
ENSMUSG00000070811	1
ENSMUSG00000034528	1
ENSMUSG00000052974	1
ENSMUSG00000074207	1
ENSMUSG00000020072	1
ENSMUSG00000020010	1
ENSMUSG00000054417	1
ENSMUSG00000052131	1
ENSMUSG00000029273	1
ENSMUSG00000037053	1
ENSMUSG00000057400	1
ENSMUSG00000030413	1
ENSMUSG00000064294	1
ENSMUSG00000010601	1
ENSMUSG00000015970	1
ENSMUSG00000025911	1
ENSMUSG00000019232	1
ENSMUSG00000021135	1
ENSMUSG00000029630	1
ENSMUSG00000029727	1
ENSMUSG00000003617	1
ENSMUSG00000040809	1
ENSMUSG00000020884	1
ENSMUSG00000025479	1
ENSMUSG00000020641	1
ENSMUSG00000025934	1
ENSMUSG00000031594	1
ENSMUSG00000058207	1
ENSMUSG00000036216	1
ENSMUSG00000033533	1
ENSMUSG00000055782	1
ENSMUSG00000028715	1
ENSMUSG00000019987	1
ENSMUSG00000025003	1
ENSMUSG00000041828	1
ENSMUSG00000087107	1
ENSMUSG00000034435	1
ENSMUSG00000036381	1
ENSMUSG00000052595	1
ENSMUSG00000030909	1
ENSMUSG00000022868	1
ENSMUSG00000021390	1

ENSMUSG00000078650	1
ENSMUSG00000021922	1
ENSMUSG00000031138	1
ENSMUSG00000029556	1
ENSMUSG00000031173	1
ENSMUSG00000000049	1
ENSMUSG00000044749	1
ENSMUSG00000028553	1
ENSMUSG00000028262	1
ENSMUSG00000030359	1
ENSMUSG00000054072	1
ENSMUSG00000048217	1
ENSMUSG00000072601	1
ENSMUSG00000027761	1
ENSMUSG00000030244	1
ENSMUSG00000040017	1
ENSMUSG00000096852	1
ENSMUSG00000021091	1
ENSMUSG00000057425	1
ENSMUSG00000039438	1
ENSMUSG00000032081	1
ENSMUSG00000070594	1
ENSMUSG00000035540	1
ENSMUSG00000025194	1
ENSMUSG00000029368	1
ENSMUSG00000020264	1
ENSMUSG00000060317	1
ENSMUSG00000021999	1
ENSMUSG00000041660	1
ENSMUSG00000003053	1
ENSMUSG00000042118	1
ENSMUSG00000032291	2
ENSMUSG00000045179	2
ENSMUSG00000004885	2
ENSMUSG00000043668	2
ENSMUSG00000021506	2
ENSMUSG00000039476	2
ENSMUSG00000021508	2
ENSMUSG00000074637	2
ENSMUSG00000079012	2
ENSMUSG00000044206	2
ENSMUSG00000025478	2
ENSMUSG00000075334	2

ENSMUSG00000038624	2
ENSMUSG00000048001	2
ENSMUSG00000086503	2
ENSMUSG00000033737	2

Table 9-3. Differentially Expressed Genes Identified from mHLCs vs E19.5.

Gene Name	Fold Change	q-value(%)
ARG1	19.47648	0
ADH1A	17.695	0
APCS	17.03517	0
GC	15.79147	0
FETUB	14.36397	0
ADH1C	14.03646	0
AZGP1	14.03505	0
CYP2C9	13.92966	0
HPR	13.09532	0
ITIH4	12.98748	0
F9	12.69693	0
HAMP	12.49579	0
ITIH3	12.13291	0
RARRES2	11.82974	0
CYP2C9	11.72359	0
CYP2C9	11.56773	0
FBP1	11.26727	0
CYP8B1	11.09845	0
KLKB1	10.8902	0
CFHR2	10.69348	0
SLCO1B1	10.69292	0
CYP2C9	10.24091	0
HAO1	10.19456	0
G6PC	10.07402	0
BAAT	9.990331	0
C4BPA	9.951422	0
F11	9.822402	0
CYP3A4	9.787597	0
LRG1	9.353939	0
SLCO1B3	8.983395	0
C3	8.711476	0
CYP3A5	8.647951	0
CRP	8.603505	0
CYP3A4	8.586851	0
SEC14L4	8.462595	0
MT1G	8.398134	0
C8A	8.368593	0
C8G	8.150217	0
AKR1C4	7.982445	0
CYP3A4	7.796432	0

AQP9	7.760693	0
AKR1C4	7.699717	0
BHMT2	7.66797	0
MBL2	7.667351	0
ETNPPL	7.637466	0
CYP2C19	7.593378	0
CYP3A4	7.584423	0
THRSP	7.560952	0
SLC2A2	7.54792	0
CPB2	7.517397	0
UGT2B15	7.491674	0
CYP4A22	7.332241	0
CYP4A11	7.332241	0
AGXT	7.318679	0
HFE2	7.294192	0
CYP2D6	6.980419	0
F12	6.977067	0
AKR1C4	6.830319	0
CYP2C19	6.818073	0
MT1M	6.80367	0
RDH16	6.790065	0
SLC27A5	6.593052	0
CYP4A22	6.47747	0
CYP4A11	6.47747	0
PAH	6.451182	0
MT1F	6.419835	0
IGFALS	6.37437	0
CYP2C19	6.368176	0
AKR1C1	6.160139	0
CYP2D6	6.125002	0
UGT2B15	5.996523	0
CYP4A11	5.921239	0
CYP4A22	5.921239	0
CYP2C19	5.913506	0
AKR1C1	5.910363	0
IGFBP1	5.869926	0
HAO2	5.860457	0
ABCG5	5.85733	0
UPB1	5.678749	0
CYP2D6	5.615751	0
LDHD	5.594874	0
CYP2D6	5.585391	0
APOA5	5.571442	0

AOX1	5.552244	0
CYP2J2	5.533126	0
MASP2	5.468354	0
NAT8	5.398458	0
SLC17A2	5.348241	0
UGT2B15	5.26995	0
AKR1C4	5.187966	0
NR1I2	5.174316	0
UGT2B15	5.170655	0
CYP4F2	5.163166	0
CYP2C8	5.15697	0
HGFAC	5.052705	0
LPA	5.042815	0
ONECUT1	5.036339	0
CYP4F12	5.031971	0
CYP2D6	4.941497	0
ABCA6	4.918435	0
AKR1C1	4.880393	0
GLS2	4.847726	0
UGT2B15	4.846635	0
HABP2	4.836134	0
UGT2B15	4.793784	0
C2	4.735973	0
AKR1C4	4.712947	0
CYP4F2	4.710266	0
TAT	4.703243	0
ABCB11	4.691866	0
VNN1	4.607717	0
F13B	4.571373	0
ONECUT2	4.554322	0
CYP2C8	4.519439	0
RTP3	4.491045	0
CYP4F12	4.488012	0
RDH16	4.441499	0
PEMT	4.402901	0
CYP2C8	4.338335	0
CSTA	4.308267	0
CD14	4.262732	0
SLC25A47	4.228052	0
ACOT12	4.156665	0
CYP39A1	4.145899	0
MLXIPL	4.130333	0
SDS	4.105943	0

PAQR9	4.014396	0

Table 9-4. Dynamically Differentially Expressed Genes Identified Between hHLCs and *In-vivo* Mouse Development.

Gene Name	spearman coeff	pearson coeff	euclidean	Foldchange human	Foldchange mouse
TPD52L1	-1	-0.99998	8.202735	3.287442	2.42286
SOX9	-1	-0.96806	6.470891	3.673472	2.691969
CA4	-1	-0.99363	12.76903	5.023663	2.934446
COL1A1	-1	-0.99799	10.52601	5.878911	1.165311
FOXM1	-1	-0.96139	6.176861	2.561048	0.105252
RIN2	-1	-0.99932	7.199617	2.972658	1.446347
LAMB1	-1	-0.98367	8.43345	3.265966	2.529975
INMT-FAM188B	-0.88264	-0.90333	8.721384	3.390595	0.190508
INMT	-0.88264	-0.90333	8.721384	3.390595	0.190508
ANGPTL2	-0.9952	-0.97192	6.669714	4.644773	1.990356
CLEC11A	-1	-0.99259	6.607506	2.787634	0.737749
MAPK13	-1	-0.99952	7.057298	2.113425	1.629894
WISP1	-1	-0.99999	7.103779	4.687063	1.852809
SPC25	-1	-0.97999	6.566221	2.608397	0.991307
SLC1A3	-1	-0.99898	6.147721	2.936563	2.013546
TYRP1	-1	-0.99885	10.5835	4.493249	0.642752
FBLN1	-1	-0.99926	9.118421	1.821108	3.43929
SCUBE2	-1	-0.98805	7.176644	3.651059	1.537498
TLE1	-1	-0.9997	6.542244	2.191705	0.648269
RAB25	-1	-0.9995	9.011451	2.633916	3.262243
ODC1	-1	-0.99905	8.152495	2.951116	0.571813
GATA5	-1	-0.95235	8.659841	3.400776	1.18373
BIRC5	-1	-0.99532	5.884918	2.165832	0.503302
MMP9	-1	-0.98558	7.253185	2.603546	0.970577
GPX3	-1	-0.99299	7.088938	3.202002	2.977199
CLDN7	-0.99789	-0.99353	7.635152	2.7394	2.987019
SPARC	-1	-0.98394	9.983665	6.048512	1.121165
P4HA2	-1	-0.99275	7.009518	2.837876	1.699379
E2F2	-1	-0.98456	6.313519	2.48639	2.249367
DPEP1	-1	-0.99968	8.169601	2.502557	2.235288
SOCS2	-0.93133	-0.90123	9.100238	2.578425	1.060604
IGF1	-1	-0.99121	6.01632	2.829677	1.577037
IGFBP3	-1	-0.99083	11.78145	7.476728	2.054755
MTHFD1	-1	-0.97191	8.034054	1.581466	2.406489
MSX2	-0.93133	-0.8851	6.676606	3.935041	2.888799
CXCL14	-1	-0.99855	6.8103	6.909237	3.966255
VCAN	-0.88264	-0.91007	6.073209	2.971706	1.744054

NID2	-1	-0.9947	9.804529	7.742699	1.517822
WNT5A	-1	-0.99715	8.458978	3.831782	2.970028
LCP1	-1	-0.99149	17.44107	5.717766	1.476532
GPC5	-1	-0.99631	7.514748	4.469395	0.45908
CTNND2	-1	-0.99652	6.886003	2.16132	2.122937
ZFPM2	-1	-0.99994	6.433396	2.97006	2.000771
ANGPT1	-1	-0.99666	7.687757	3.271078	2.445031
CDH10	-1	-0.99725	9.986773	5.275753	1.360325
COL2A1	-1	-0.97266	6.112505	2.43766	4.138637
PPP1R1A	-1	-0.99688	6.177531	3.719291	2.619232
DGAT1	-1	-0.99811	7.963457	2.405352	0.449161
ARC	-1	-0.98125	6.23039	0.718595	2.144159
SNAI2	-1	-0.99993	11.54102	5.39566	3.316863
PROS1	-1	-0.97454	9.90243	3.815411	1.108385
CLIC6	-1	-0.99905	7.546699	3.349008	2.775784
RCAN1	-1	-0.99644	6.985294	2.497737	1.162306
SMOC2	-1	-0.97742	5.756148	3.828843	1.806198
CBS	-1	-1	6.806959	3.699452	2.540984
EPB41L3	-1	-0.99894	6.68763	2.424304	2.302699
ANKS1A	-1	-0.99574	5.626824	3.292288	0.491621
BAMBI	-0.94094	-0.91258	7.815337	2.692172	1.648303
SYT4	-1	-0.99943	7.897645	4.126923	1.209736
BIN1	-1	-0.99905	5.915923	2.698904	1.554763
SNCAIP	-1	-0.99984	7.299104	4.133323	1.46805
PPIC	-1	-0.98817	6.57819	2.855431	1.555656
ASRGL1	-1	-0.9768	5.948578	2.782423	1.639415
SLC1A1	-1	-0.97201	5.639671	2.737276	0.604882
HSPA12A	-1	-0.98903	6.379916	2.270736	1.029907
BNC1	-1	-0.99934	5.736315	3.709759	1.794152
PSTPIP2	-1	-0.98109	6.241606	2.436976	0.941142
TK1	-1	-0.99993	6.948543	2.809313	1.23098
COL3A1	-1	-0.97662	16.61379	7.134201	1.788591
IGFBP5	-1	-0.98632	8.864169	4.590883	4.055705
RGS1	-1	-0.98933	5.666029	2.867679	0.460585
RGS2	-0.94958	-0.95763	7.162765	3.418134	1.389152
KIFAP3	-1	-0.99948	7.615188	2.139961	2.056832
FRZB	-1	-0.99192	10.56679	3.121508	2.94741
PAMR1	-0.93133	-0.91952	5.721357	5.799499	3.246058
MDK	-1	-0.99438	7.171142	1.836701	3.296871
SPINT1	-1	-0.99904	8.263304	2.241655	2.977843
BMP2	-1	-0.99857	9.248287	2.995835	0.169777
CD93	-1	-0.99244	7.169902	2.578764	1.234445
PKIA	-1	-0.99594	7.832302	3.785553	2.247848

POSTN	-1	-0.98873	9.225436	4.444603	3.342483
SERPINI1	-1	-1	5.693185	2.218662	1.623499
FAM198B	-1	-0.98209	9.50968	3.783681	2.424531
CCDC109B	-1	-0.99142	7.50031	2.350011	2.384037
TDO2	-1	-0.99771	8.91894	4.859247	1.872948
PDGFC	-1	-0.99708	7.619929	3.268998	3.122205
PITX2	-1	-0.98387	11.57107	4.543715	3.106901
F3	-1	-0.99305	7.140429	3.964018	2.137135
COL15A1	-1	-0.99993	10.7584	9.604504	1.286437
ALDH4A1	-1	-0.99651	7.549047	1.174329	2.871219
SEMA3C	-1	-0.99344	7.557414	2.411146	2.100976
PCDH7	-1	-0.99805	6.056629	2.391598	1.863397
Una.	-1	-0.99988	5.914584	3.745275	2.438067
CDX2	-1	-0.9971	7.904824	3.482707	2.92062
COL1A2	-1	-0.98409	11.00918	6.08499	1.52914
AASS	-1	-0.99866	8.450191	1.215457	3.420273
ADD2	-1	-0.99979	6.627112	1.916887	2.374459
SLCO2B1	-1	-0.98626	11.4765	5.186084	1.200618
TSC22D3	-1	-0.9959	8.455545	3.045552	1.354535
GPR124	-1	-0.984	6.75361	5.63952	2.42959
VEGFC	-1	-0.99946	8.449825	4.848861	2.020543
PLAT	-1	-0.99529	7.554406	6.14278	3.220333
SLIT2	-1	-0.97949	9.463614	3.435444	2.839977
SCRG1	-1	-0.9995	11.93331	5.435314	1.028461
KLKB1	-1	-0.99457	9.584776	3.622454	3.756616
CDH11	-0.85143	-0.90562	10.08777	5.175973	2.01917
MT1X	-1	-0.99958	14.69135	3.97352	5.361478
MT1M	-1	-0.99181	10.15712	4.699498	4.710045
MT1E	-1	-0.99001	11.40912	3.520225	4.581416
MT1G	-1	-0.99984	12.87808	3.050806	4.880415
MT1F	-1	-0.99975	11.62271	3.417577	4.523986
MT1H	-1	-0.99544	9.078276	2.501644	4.216505
CRISPLD2	-1	-0.99808	9.640576	5.248432	4.158231
CDH13	-1	-0.9912	5.85786	5.24851	0.446286
PDGFD	-1	-0.99941	5.637534	2.570399	1.776597
THY1	-1	-0.9998	5.898586	2.943447	1.972785
CADM1	-1	-0.99126	5.666782	2.323375	2.465708
APOA4	-1	-0.9905	19.40484	7.492938	2.460919
MNS1	-1	-0.98452	7.064168	2.397407	0.835902
DNAJA4	-1	-0.99992	6.344101	0.936862	2.376461
TSPAN3	-1	-0.99951	6.062608	1.538401	2.214344
LOXL1	-0.99673	-0.97118	6.863864	3.662788	1.71054
PLOD2	-1	-0.99999	6.69004	4.37611	2.075458

IGDCC3	-0.92077	-0.87152	8.968043	3.941408	3.504619
SLCO2A1	-1	-0.99894	6.527005	4.086047	2.784743
SBSPON	-1	-0.92368	6.908534	5.484962	1.188689
RASGRP2	-1	-0.9936	6.149767	0.681514	2.034262
CST1	-1	-0.99627	7.394144	8.047238	1.132582
PRSS35	-0.99328	-0.97129	11.29819	4.156408	2.396574
UCP2	-1	-0.99942	8.215212	1.930647	2.590443
DIO1	-1	-0.96993	14.54904	6.130609	0.679016
PARM1	-1	-0.9977	12.20437	4.260447	2.812948
KANK4	-1	-1	9.194998	4.579135	1.766136
FOXA1	-1	-0.99185	8.920911	4.634631	1.352117
TGFBI	-1	-0.99385	13.99386	6.888526	4.47516
SLC51A	-1	-0.99976	11.17642	4.114251	2.0838
PRTG	-1	-0.99459	7.032647	2.896586	3.573539
SOSTDC1	-1	-0.96966	7.064512	2.513585	2.79548
ALX1	-1	-0.99942	8.59478	1.798566	2.726357
NODAL	-1	-0.99778	5.878139	3.996051	1.314737
HAND1	-1	-0.99924	14.10207	6.999903	4.126651
CDKN1C	-1	-0.96259	11.07485	3.927051	1.347614
CPE	-1	-0.99788	6.760178	2.924181	1.529346
SPON1	-1	-0.94975	7.059034	3.713748	1.894611
SLC19A3	-1	-0.99826	5.850146	2.924093	0.624851
RGS4	-0.81513	-0.90034	6.579283	4.283272	1.601526
LEFTY2	-1	-0.99985	10.5707	7.257231	3.448976
LEFTY1	-1	-0.99918	12.09782	6.497584	3.566699
ECHDC3	-1	-0.99237	6.323923	1.652374	2.591008
TGFB2	-1	-0.99998	5.715218	2.724041	2.147273
PRSS23	-1	-1	7.360117	5.920373	1.979309
CITED2	-1	-0.99871	5.875684	1.816708	2.169233
GPR176	-1	-0.9713	7.38363	1.459685	3.040128
NPNT	-1	-0.99812	8.876435	4.837207	1.820217
FBXO2	-1	-0.9987	5.900228	3.106349	2.368491
AMOT	-1	-0.97965	9.637363	4.032142	2.364655
KCNJ2	-1	-0.98704	6.847413	3.717141	1.297968
S100A10	-1	-0.99657	8.964101	3.160094	1.51785
SLC35F2	-1	-0.98804	6.085977	2.16187	2.704036
MFAP4	-1	-0.99958	9.761389	7.621643	1.672373
MUM1L1	-1	-0.99995	11.21239	3.725983	2.428555
CXXC4	-1	-0.99765	7.163809	2.69722	0.993399
SHISA2	-1	-0.96374	5.762676	4.010679	3.477351
TMEM88	-1	-0.9962	8.281414	4.204806	1.685703
EPCAM	-1	-0.99605	8.012674	1.421146	3.699162
TSHZ1	-1	-0.99768	5.76684	3.549243	2.490843

OSR1	-1	-0.99864	5.911354	4.89944	3.034315
NDNF	-1	-0.99573	7.755868	4.959952	2.273667
FAM110B	-0.99971	-0.9615	5.893586	2.858926	2.0205
TMEM200A	-1	-0.98677	6.609809	4.54988	2.471675
FLRT3	-1	-0.99992	11.71001	2.857839	3.323295
RSPO2	-1	-0.98042	5.719607	4.603523	2.394453
REEP1	-1	-0.99862	9.033056	4.04275	0.852762
CTHRC1	-1	-0.99862	8.530364	4.270829	2.735034
TMEM119	-1	-0.99976	7.568417	5.020092	3.16296
HBZ	-1	-0.99451	7.102326	2.841429	4.268472
SLIT3	-1	-0.9889	6.504	3.136287	2.459745
CD248	-1	-0.97829	7.223254	6.151601	2.975828
MAB21L2	-1	-0.99982	8.168393	2.870649	2.087563
GPC6	-1	-0.99985	6.087021	3.077187	1.868578
DSC3	-1	-0.98015	7.702813	2.949513	1.052887
IFITM1	-1	-0.98948	6.026868	4.574779	1.737948
CDH3	-1	-0.99443	6.323749	3.145868	2.18805
MYL4	-1	-0.98138	7.870759	2.461171	3.26288
AKR1B10	-1	-0.99588	7.158083	2.770126	1.387845
HS6ST2	-1	-0.99254	6.006891	3.113056	2.3427
CYTL1	-1	-0.99661	6.758688	3.667129	0.915346
CDC42EP5	-1	-0.99727	6.718535	6.34062	1.102656
IFITM1	-1	-0.99015	6.506348	4.55697	1.63668
S1PR3	-0.95736	-0.93724	6.881802	2.458088	2.214965
TMEM100	-1	-0.99171	5.963785	3.954433	2.495345
CLDN3	-1	-0.98888	8.725134	2.710618	1.899637
THBD	-1	-0.99986	11.04372	3.036152	0.803528
SERPINA3	-0.9805	-0.94329	5.634335	3.146055	2.260811
TNNC1	-1	-0.99895	5.767938	3.582013	3.104565
KCTD12	-0.94094	-0.93558	8.172536	3.826093	2.355453
TGFBR3	-0.92077	-0.89025	5.029718	6.453567	1.868401

Table 9-5. Identified Dynamically Differentially Expressed Genes that Retain the Same Dynamics in hHLCs and mHLCs

Gene Name
ADD2
ALDH4A1
ANGPT1
ANKS1A
APOA4
ARC
BIN1
BIRC5
BNC1
CCDC109B
CDC42EP5
CDH11
CDH13
CDH3
CDX2
CITED2
CLEC11A
COL15A1
COL1A1
COL3A1
CPE
CRISPLD2
CXCL14
CXXC4
DGAT1
DIO1
DPEP1
DSC3
E2F2
ECHDC3
FAM110B
FAM198B
FBLN1
FBXO2
FLRT3
FOXM1
GATA5
GPX3
HAND1

HSPA12A
IGF1
INMT
KANK4
LAMB1
LCP1
LEFTY1
LEFTY2
MAPK13
MDK
MNS1
MTHFD1
MYL4
NDNF
NODAL
OSR1
P4HA2
PARM1
PDGFC
PDGFD
PKIA
PLAT
PROS1
RGS1
RGS4
RIN2
RSPO2
S1PR3
SBSPON
SCUBE2
SEMA3C
SERPINA3
SERPINI1
SHISA2
SLC1A1
SLC1A3
SLC51A
SLCO2B1
SLIT2
SNAI2
SOCS2
SOSTDC1
SOX9

SPARC
SPC25
SPINT1
SPON1
SYT4
THY1
TK1
TMEM100
TMEM119
TNNC1
TSC22D3
TSHZ1
TSPAN3
UCP2
VCAN
VEGFC
WNT5A